

中国科学技术大学

专业硕士学位论文

(专业学位类型)



SMT 贴片元件自动光学检测关键技术 研究

作者姓名: 桂旺友
专业领域: 计算机技术
校内导师: 赵云波 教授
实践导师: 许镇义 副研究员
完成时间: 二〇二四年五月二十四日

University of Science and Technology of China
A dissertation for master's degree

(Professional degree type)



**Research on Key Technologies for
Automated Optical Inspection of
SMT Chip Components**

Author: Gui Wangyou

Speciality: Computer Technology

Supervisor: Prof. Zhao Yunbo

Advisor: AR. Xu Zhenyi

Finished time: May 24, 2024

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名： 程旺友

签字日期： 2024年5月24日

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

控阅的学位论文在解密后也遵守此规定。

公开 控阅（___年）

作者签名： 程旺友

导师签名： 高之洪

签字日期： 2024年5月24日

签字日期： 2024年5月24日

摘 要

近年来，表面贴装技术（Surface Mounted Technology, SMT）广泛地应用于电子组装行业，高效的生产流程极大地促进了电子产品的智能化、精密化发展。印刷电路板（Printed Circuit Board, PCB）上的布线和元件随之变得越来越密集，这对其在生产过程中的质量检测问题也提出了更大的挑战。目前主流的方法是采用自动光学检测（Automated Optical Inspection, AOI）技术对贴片元件经过回流焊后的可能产生的缺陷情况进行检测。然而，当前的 AOI 技术大多还采用基于模板匹配的传统视觉手段，在实际生产中仍然存在着较多的误检和漏检情况，且整体工作效率不高，因此对其进一步地研究和改进有着重要的意义。

本文针对实际产线上的 AOI 系统的生产情况进行分析。研究发现，AOI 系统的弊端主要在于使用的技术缺乏先进性，导致效率低下。一方面，由多张 PCB 局部图像拼接得到的整板图像质量不高，存在着明显的错位情况，这严重影响着有缺陷区域的识别效果；另一方面，在缺陷检测时使用的模板匹配技术存在着较大的弊端，该算法高度依赖于模板和拼接图像的质量，前者需要专业人员精心设计建模，后者又取决于图像拼接的准确度。另外，匹配时的判定规则也比较单一，难以应对复杂的缺陷特征，使得误检率达到了 80%，以至于需要大量人工复检工作。针对以上问题，本文对这两个环节中所使用的关键技术进行研究，对图像拼接算法进行了改进，并提出一种全新的缺陷检测算法，具体内容包括：

1. 针对 AOI 系统采集的若干张 PCB 局部图像拼接成 PCB 整板图像的问题，将硬拼接和软拼接的方式相结合，提出一种基于 SURF 主要距离的主板图像拼接算法，实现了 PCB 图像的实时高质量拼接。针对传统的 SURF 拼接算法进行改进，提出基于划分区域的特征点提取方法和基于主要范围的匹配对细筛选方法，在算法速度和准确度上都得到了较好的提升。以硬拼接流程为基础，为应对由于震动导致的不同组相邻图像间偏移量一直在变化的情况，利用改进后的 SURF 算法提取出待拼接图像的特征点并匹配，将所有匹配对之间主要距离作为偏移量，在每次拼接不同对象时都能做到实时调整。为了弥补光照不均造成的拼接缝两边色差明显，采用基于加权像素的空域融合法处理该区域。以最先完成拼接的第一列图像为准，通过偏移补偿方法保证所有其他列拼接图像尺寸与之一致。在 PCB 整板图像拼接实验中，本文方法只需要 16s 左右的时间就可以在保留图像高清晰度的前提下稳定地实现图像拼接过程，并且最终结果图的错位情况得到明显改善。

2. 针对 SMT 产线中贴片元件经过回流焊后的缺陷检测问题，结合深度学习技术，将目标检测应用于贴片元件的缺陷检测任务，提出一种基于自适应协同特

征的 DETR 缺陷检测算法，实现了多种复杂缺陷元件的有效检测。为应对相关数据稀少的情况，构建了 SMT 贴片元件缺陷数据集，包含多种存在不同缺陷情况下的 PCB 图像。考虑到工业场景中对检测速度和精度的高要求，采用轻量化的 RT-DETR 作为基准模型，在借助 transformer 强大的建模能力的同时，还具有较好的实时性。针对贴片元件特殊的缺陷形态，设计出自适应协同模块替代部分传统的卷积模块，能够更好地关注到元件缺陷的形变特征。提出一种全新的损失函数，对 PCB 图像中的微小贴片元件更加友好，并且具备较快的收敛速度。在自行收集贴片元件缺陷数据集的实验表明，本文方法在 mAP50 指标上达到了 88.3%，相对于基准方法提升了 8.4%，在 FPS 指标上达到了 90，在速度和精度上达到了较好的平衡，优于目前所有的实时目标检测模型。

基于上述研究，本文实现了 SMT 贴片元件自动光学检测关键技术上的改进，这些算法均将部署在课题组搭建的 AOI 设备上，预计建成后将有效提升制造商的整体生产效率，并产生巨大经济效益。

关键词：表面贴装技术；自动光学检测；图像拼接；缺陷检测

ABSTRACT

In recent years, Surface Mounted Technology (SMT) has been widely used in the electronics assembly industry. Its efficient production process has greatly promoted the intelligent and precision development of electronic products. Consequently, the wiring and components on Printed Circuit Boards (PCB) have become increasingly dense, posing greater challenges for quality inspection during the production process. Currently, the mainstream method involves using Automated Optical Inspection (AOI) technology to detect possible defects in surface-mounted components after reflow soldering. However, most current AOI technologies still rely on traditional vision methods based on template matching, leading to a significant number of false positives and false negatives during actual production, and overall low efficiency. Therefore, further research and improvement of AOI technology are of significant importance.

This dissertation analyzes the production situation of AOI systems on actual assembly lines. The research findings indicate that the main drawbacks of AOI systems are primarily linked to the use of outdated technology, leading to low efficiency. On one hand, the overall image derived from stitching multiple partial PCB images is of poor quality, displaying noticeable misalignments that significantly impact the recognition of defective areas. On the other hand, the template matching technique used for defect detection has significant drawbacks. This algorithm heavily relies on the quality of the template and the stitched image. The former requires meticulous design and modeling by professional personnel, while the latter depends on the accuracy of image stitching. Moreover, the decision-making rules during matching are relatively simplistic, making it challenging to address complex defect features, resulting in an 80% false positive rate, necessitating significant manual rechecking work. To address these issues, the dissertation conducts research on the key technologies utilized in these two processes, enhances the image stitching algorithm, and proposes a novel defect detection algorithm. Specific details include:

1. Aiming at the problem of stitching several PCB local images acquired by the AOI system into PCB whole-board images, combining the hard stitching and soft stitching methods, we propose a motherboard image stitching algorithm based on the main distance of the SURF to realize real-time high-quality stitching of PCB images. Improvements have been made to the SURF algorithm by proposing a feature point extraction method based on partitioning regions and a matching pair fine screening method

based on the primary scope, resulting in a certain enhancement in speed and accuracy. Based on the hard stitching process, in order to cope with the situation that the offsets between different neighboring images are always changing due to vibration, the improved SURF algorithm is used to extract the feature points of the images to be spliced and match them, and the main distance between all matched pairs is taken as the offset, which can be adjusted in real time each time when stitching different objects. In order to compensate for the obvious color difference between the two sides of the stitching seam caused by uneven illumination, weighted pixel-based null space fusion is used to deal with this region. The first column of the spliced image is the first to be completed, and the offset compensation method ensures that all the other columns of the spliced image size is consistent with it. In the experimental PCB image stitching, the method proposed in this dissertation only requires approximately 16 seconds to stably achieve image stitching while maintaining image clarity, leading to a noticeable improvement in the misalignment of the final stitched image.

2. Aiming at the problem of defect detection of SMT chip components after re-flow in SMT production lines, combined with deep learning technology, object detection method is applied to the defect detection task of SMT chip components, and a DETR defect detection algorithm based on adaptive collaborative features is proposed, which realizes the effective detection of multiple complex defective components. In order to cope with the scarcity of relevant data, a SMT chip component defect dataset is constructed, containing a variety of PCB images with different defects. Considering the high requirements for detection speed and accuracy in industrial scenarios, the lightweight RT-DETR is adopted as the benchmark model, which has good real-time performance while leveraging the powerful modeling capability of transformer. Aiming at the special defect morphology of chip components, an adaptive cooperative module is designed to replace part of the traditional convolution module, which can better focus on the deformation characteristics of component defects. A new loss function is proposed, which is more friendly to the tiny chip components in PCB images and has a faster convergence speed. The experimental results from our self-collected dataset of surface-mounted component defects indicate that our method achieved an mAP50 score of 88.3%, showing an improvement of 8.4% over the baseline method. It also achieved an FPS score of 90, demonstrating a good balance between speed and accuracy, thereby outperforming all current real-time object detection models.

Based on the above research, this dissertation realizes the improvement on the key technology of SMT chip component defect detection, and these algorithms will be de-

ployed on the AOI equipment built by the group, which is expected to effectively improve the overall production efficiency of the manufacturer and generate great economic benefits after the completion of the project.

Key Words: Surface Mounted Technology (SMT); Automated Optical Inspection (AOI); Image Stitch; Defect Detection

目 录

第 1 章 绪论	1
1.1 研究背景与意义	1
1.2 国内外研究现状	2
1.2.1 自动光学检测技术研究现状	2
1.2.2 PCB 图像拼接技术研究现状	4
1.2.3 PCB 缺陷检测技术研究现状	4
1.3 研究内容	6
1.3.1 基于 SURF 主要距离的主板图像拼接算法研究	6
1.3.2 基于自适应协同特征的 DETR 缺陷检测算法研究	6
1.4 本文的组织结构	7
第 2 章 相关技术	9
2.1 AOI 系统	9
2.1.1 组成结构	9
2.1.2 工作原理	11
2.1.3 关键技术分析	13
2.2 图像拼接理论基础	13
2.2.1 图像滤波	13
2.2.2 SURF 特征检测算法	15
2.3 缺陷检测理论基础	19
2.3.1 基于卷积神经网络的目标检测算法	19
2.3.2 基于 transformer 的目标检测算法	21
第 3 章 基于 SURF 主要距离的主板图像拼接算法	23
3.1 引言	23
3.2 图像采集	23
3.3 图像预处理	24
3.3.1 实采图像分析	24
3.3.2 图像去噪	24
3.4 图像拼接算法设计	25
3.4.1 图像拼接流程	26
3.4.2 硬拼接	26
3.4.3 基于 SURF 改进的软拼接	27

3.4.4 基于 SURF 主要距离的主板图像拼接算法	32
3.5 实验设计与分析	34
3.5.1 实验数据	34
3.5.2 实验环境	35
3.5.3 评价指标	35
3.5.4 实验结果分析	37
3.6 本章小结	39
第 4 章 基于自适应协同特征的 DETR 缺陷检测算法研究	41
4.1 引言	41
4.2 RT-DETR	42
4.3 改进的 RT-DETR 模型设计	44
4.3.1 自适应协同模块	44
4.3.2 损失函数	46
4.4 实验设计与分析	49
4.4.1 实验数据	49
4.4.2 实验环境	54
4.4.3 评价指标	54
4.4.4 实验结果分析	55
4.5 本章小结	60
第 5 章 总结与展望	61
5.1 现阶段工作总结	61
5.2 未来工作展望	62
参考文献	63
致谢	68
在读期间发表的学术论文与取得的研究成果	69

插图清单

图 1.1	SMT 生产流程	2
图 2.1	AOI 系统结构图	10
图 2.2	AOI 系统实物图	10
图 2.3	S 形采图路线	11
图 2.4	RGB 光源	12
图 2.5	AOI 系统工作流程图	12
图 2.6	生成高斯模板	14
图 2.7	整数型高斯模板	15
图 2.8	积分图像	16
图 2.9	盒式滤波器	17
图 2.10	盒式滤波	17
图 2.11	特征点主方向	18
图 2.12	特征点描述子	18
图 2.13	R-CNN 算法流程图	20
图 2.14	YOLO 算法流程图	21
图 2.15	DETR 算法流程图	22
图 3.1	行偏移和列偏移	24
图 3.2	图像预处理效果	25
图 3.3	PCB 图像拼接流程	26
图 3.4	相邻图拼接流程	27
图 3.5	行拼接区域划分	28
图 3.6	SURF 特征点	29
图 3.7	KNN 匹配结果	30
图 3.8	匹配点距离分布	31
图 3.9	基于主要范围的筛选法处理结果	31
图 3.10	SURF 拼接图像	32
图 3.11	主要匹配点距离分布	33
图 3.12	拼接实验数据部分示例	35
图 3.13	主板 1 拼接成像图	39
图 4.1	SMT-DETR 结构图	44

图 4.2	OrthoAttention 示意图	45
图 4.3	DCN 采样示意图	45
图 4.4	正交注意力模块结构图 (S5、S6)	46
图 4.5	自适应协同模块结构图 (S7)	46
图 4.6	SMTC 数据集部分示例	50
图 4.7	AOI 系统实际采集数据	50
图 4.8	单个元件缺陷图	51
图 4.9	单一元件图 RoI 提取	52
图 4.10	PasteChips 两种方法生成的图像	53
图 4.11	DeepPCB 数据集部分示例	53
图 4.12	VisDrone2019 数据集部分示例	54
图 4.13	不同模型在相同的 SMTC-Origin 数据集测试样本上的检测结果	57
图 4.14	模型训练时 mAP 曲线. (a) mAP50; (b) mAP50-95	59
图 4.15	模型训练时 loss 曲线 (a) L_{box} . (b) L_{cls}	59

表格清单

表 2.1	工控机配置参数	9
表 2.2	AOI 光源主要参数	11
表 3.1	行偏移和列偏移测量	27
表 3.2	拼接实验环境配置参数	35
表 3.3	相邻图像拼接对比实验结果	37
表 3.4	相邻图像拼接消融实验结果	38
表 3.5	PCB 整板图像整体拼接实验结果	39
表 4.1	检测实验环境配置参数	54
表 4.2	混淆矩阵	55
表 4.3	模型训练时部分超参数设置	56
表 4.4	SMTC-Origin 数据集对比实验结果	56
表 4.5	DeepPCB 数据集对比实验结果	57
表 4.6	VisDrone-DET2019 数据集对比实验结果	57
表 4.7	改进模块消融实验	58
表 4.8	将 ACBlock 插入到主干网络的不同位置的结果。{S5,S6,S7} 表示同时插入到 S5,S6,S7 三个位置, {S5} 表示只插入到 S5, 其他同理。	58
表 4.9	SMTC-Pseudo 实验结果	58

第1章 绪 论

本章内容主要介绍本文工作的研究背景和意义，相关技术的研究现状，以及接下来的行文安排。

1.1 研究背景与意义

科技的进步带动生产力的发展，当今时代，智能制造已经成为我国制造业的发展重心。2021年，政府发布的《“十四五”智能制造发展规划》中也明确提到智能制造为主攻方向的重要地位。表面贴装技术（Surface Mounted Technology, SMT）是当前电子制造行业中应用最为广泛的组装技术，也是智能制造技术升级的重点关注对象。目前我国的PCB制造业规模居世界第一，但在先进性方面与国外仍存在差距。截止到2022年，我国的SMT设备市场大约占有全球市场的39%份额，之后是欧洲和北美地区，分别占有14%和11%的市场份额。然而，SMT设备的主要厂商却分布在欧美与日韩地区，他们占据了近50%的生产额，造成了我国进口与出口的极大不平衡。因此，SMT工艺技术水平的高低将直接影响着我国电子制造行业的前景^[1]，推进SMT工艺技术的优化发展有着重要的意义和价值。

SMT是一种直接将无引线或短引线的片式元件或器件贴装到印制电路板（Printed Circuit Board, PCB）表面规定的位置上，在无源基板上进行高密度组装各种片式元件的表面贴装技术^[2]。随着电子产品趋向于多功能化、智能化、小型化^[3]，SMT取代了以往落后的插件封装技术，极大地提高了PCB组装（Printed Circuit Board Assembly, PCBA）的生产效率。然而，随之而来的，SMT也给生产过程中质量检测问题带来了更高的挑战。完整的SMT生产流程如图1.1所示，其中缺陷检测环节是就把控电子产品质量的关键，目的在于对回流炉焊接后的贴片元件可能产生的缺陷情况进行检测，以避免生产的次品进入后续流程。随着印刷电路板上线路图形精细化、片式元件微型化及SMT组件高密度组装、快速组装的发展趋势，自动光学检测（Automated Optic Inspection, AOI）已经代替了传统的人工目检成为当下最主流的检测方式^[3]。

目前应用最为广泛的AOI技术是一种基于光学原理和机器视觉的手段。具体来说，通过工业相机多次拍摄PCB表面图像并拼接成完整大图，再与标准模板对比找出所有需要检测的区域，最后通过识别算法确定当前区域的元件是否有缺陷。尽管这在很大程度上提升了检测效率，但随着技术的升级，当前的AOI技术在面对元器件集成度日益增加的PCBA时，也显得疲于应付。由于贴片元

件微小且排列密度较高、拍摄图像容易受到环境干扰、图像拼接过程需要消耗大量时间等因素的影响，当前的 AOI 技术表现出响应速度慢、缺陷误检率高等问题，整体效率低下，且需要人工复判。据某知名消费电子制造商实际产线的数据可知，因此，对当前的 AOI 技术进行进一步地研究和升级有着重要的意义。



图 1.1 SMT 生产流程

本文通过深入调研，总结出两点引起目前存在问题的原因。一方面是图像拼接的问题，在拍摄 PCB 图像时，由于相机视场和图像分辨率的双重要求，需要拍摄多张局部图像以拼接成完整图像，受到 PCB 震动的影响，相邻图像间存在着不同程度的平移和旋转误差，且采集的图像容易受到环境影响，使得拼接后的图像存在着明显的错位情况；另一方面是缺陷检测的问题，基于模板匹配的方式仍然存在着较大的弊端，该算法高度依赖于匹配双方的质量，即模板和拼接图像，前者需要专业人员精心设计建模，后者又取决于图像拼接的准确度。另外，匹配时的判定规则也比较单一，难以应对复杂的缺陷特征，使得误检率较高。

本文积极响应国家高水平发展制造业自动化和智能化的号召，紧抓当前 AOI 技术的痛点，针对其中的关键技术进行研究，实现了更加稳定高效的图像拼接算法和轻量化的贴片元件缺陷检测算法，对于提升我国电子科技制造质量检测水准有着积极意义。

1.2 国内外研究现状

1.2.1 自动光学检测技术研究现状

早在 20 世纪 70 年代，以美、日、韩为代表的半导体行业强国就开始着眼于利用自动化检测替代电子产品生产中的人工目检，不过由于此时的制造业水平

整体较为落后，质检的目的大多在于检测元件重复性较高的电子产品^[4]。进入80年代，SMT技术掀起了电子组装行业技术革命的浪潮，传统的大体积插件被替代为微小的贴片元件，对元件贴装时的好坏进行检测就显得格外重要。随着机器视觉的发展，将光学检测应用于SMT流程中成为了必然，AOI技术应运而生。

早期的AOI设备主要由当时的制造业强盛的国际化大公司开发，例如日本的Omron、韩国的Koh Young、美国的Agilent、德国的Viscom AG等，几乎垄断了世界的AOI技术。我国的AOI技术应用始于20世纪80年代，但刚开始处于空白期，无论是贴片机还是AOI设备都完全依赖于进口。高昂的设备价格使得不少电子行业的公司望而却步，直到2003年，神州视觉成功研发出了首台完全具备自主知识产权的国产AOI检测设备ALeader，这一里程碑事件将国内的AOI技术研发推上了日程，吸引越来越多的专业人士投身于AOI技术的研究与开发事业^[5]。到2005年，国内已经有了不少AOI设备厂商，他们大多分布于沿海经济发达地区，其中有代表性除神州视觉外还有福心光电、振华兴科技有限公司等等。2010年以后，我国的AOI技术进入了快速发展期。再到如今，随着AOI技术不断创新升级，目前我国的AOI设备主要以2D AOI为主，国产化率已经达到了100%，3D AOI仍在起步阶段。

从AOI的软硬件系统设计，以及在检测缺陷时的分析和判别规则来看，目前主要有两种检测方法，分别是设计规则检测法和图形识别法^[6]：

(1) 设计规则检测用于验证电路板制造过程中是否符合设计规范和设计要求。AOI系统先采用图像处理算法来提取元件位置、焊接质量、线路连接等设计规则所需的特征信息，然后利用机器学习算法或预先定义的规则，对提取的特征进行分析和比对，以检测元件位置、焊接质量、线路连接等是否符合设计规则要求。设计规则检测对检测的正确性有着较好的保证，但因为是基于预先定义的规则和模式进行检测和比对，因此可能对非标准或定制化的产品缺乏灵活性，且制定规则需要专业的知识和经验。

(2) 图形识别法利用机器视觉的技术对采集的图像进行分析，首先对图像进行特征提取，例如边缘检测、形状分析、颜色识别等，以捕获电子设备中的各种元件、线路和焊接等特征，然后通过识别算法，如模式匹配算法，将提取的特征与预先定义的图形模式或设计规则进行比对，以识别元件的位置、方向、大小等，以及线路的连接状态和焊接质量^[7]。目前该方法已成为AOI检测的主流方法，并且随着更多先进的图像识别技术的应用，使得检测性能不断提升，另外可视化后的检测结果也有利于后续的反馈工作。不足之处在于该方法需要高质量的图像数据，生产环境中的光照、机械振动等因素可能会影响图像的质量，从而影响识别结果的准确性。

1.2.2 PCB 图像拼接技术研究现状

图像拼接技术旨在将两幅或者多幅具有重叠区域的图像无缝合并为一幅图像，在计算机视觉领域有着广泛的应用^[8]。一般来说，主要有两种分支：基于像素到像素的直接方法和基于特征的拼接算法^[9]。后者在不同的应用场景中具有更强的鲁棒性，因此成为当下研究的热点，其主要实现步骤包括：

(1) 特征点提取和匹配。这一步骤旨在寻找和匹配不同图像中的共同特征点。常用的特征检测算法包括尺度不变特征变换 (SIFT)^[10-11]、加速鲁棒特征 (SURF)^[12]和方向梯度直方图 (HOG)^[13]。

(2) 图像配准。在匹配特征点后，需要对图像进行配准，以消除图像之间的几何失真。常用的配准算法包括单应性变换、仿射变换和透视变换。

(3) 图像融合。配准后的两幅图像已经建立好统一坐标系，需要图像融合在一起，消除多余的重叠部分。通常采用像素级的融合算法，它包括基于空域和基于变换域的融合。

PCB 图像拼接的目的在于将 AOI 系统采集的一系列相邻间有重叠的局部图像，组合成一张完整的 PCB 整板图像。由于在图像采集过程中容易受到机械振动、光照等环境影响，对拼接算法的稳定性要求非常高。近年来，随着诸多学者的深入研究，PCB 图像拼接技术取得了显著进展。Li 等^[14]通过使用双 Mark 点和 Mark 圈内的快速区域匹配，有效避免了噪声干扰或噪声屏蔽特征点的对准误差，提高了图像拼接的准确性。裴新权等^[15]开展的研究专注于开发一种专门针对 PCB 裸板检测的图像拼接技术。首先，该技术运用 SURF 算法提取相关特征，接着采取一种针对 PCB 缺陷特征改进的加权合并策略以完成图像的合并。根据实验的结果，该技术能够有效地进行 PCB 裸板图像的无缝合并处理。然而，对于 PCB 焊接图像的合并处理，则尚存在技术难题。杨帆等^[16]通过在 PCB 板上设置标记，并依据标记点的固定匹配关系来剔除无效的 SIFT 特征点匹配对，可以很好地应对具有图像中具有多个可重复特征元素的情况，但只针对 PCB 裸板拼接。

1.2.3 PCB 缺陷检测技术研究现状

SMT 场景中的贴片元件缺陷检测一直以来都是困扰着工业界的问题。不仅由于缺陷本身的复杂性导致检测困难，而且工业场景中对实时性的要求也非常严苛。另外，由于工业缺陷数据通常非常难收集，至今为止没有相关的贴片元件缺陷公开数据集供人们研究，大大阻碍了技术的发展。本节简要概述了目前工业界使用的主要的两种缺陷检测方法，包括传统图像处理方法和深度学习方法。

目前工业界常用的一种方法是基于模板比对的 AOI 技术。通过将产线上拍摄的多张 PCBA 图像拼接完成后与标准模板进行比对，定位出待检测区域，并通

过判断准则识别缺陷。然而，该方法严重依赖于各个流程中子任务的完成度，主要在于模板的制作、拼接图像以及判定准则的准确度。通常情况下，模板需要专业工程师进行制作、图像拼接过程需要兼具速度和准确度，而且机械的判定准则无法确定灵活的缺陷边界，这就不可避免地存在着严重的误检问题，因此通常需要在机器检测后进行大量的人工复判。考虑到模板匹配算法的不足，不少学者对此进行了研究。Sundaraj 等^[17]使用颜色背景减法来检测贴片元件的缺件和偏移。Cai 等^[18]首次从目标检测的角度切入，采用视觉方法提取背景中的有效特征来检测缺陷，具有较强的通用性。Zhang 等^[19]结合颜色直方图和 SIFI 特征，提出一种两阶段级联放大（CZI）的焊点检测方法，以应对实际情况大部分元件都是无缺陷的情况。

深度学习的快速发展为 SMT 贴片元件缺陷检测提供了新的解决方案，然而针对 PCBA 上贴片元件缺陷的检测研究较少，目前主要的研究也只是针对于提取待检测区域后的缺陷分类工作。如，YKim^[20]使用垂直和水平投影提取 PCB 图像中的贴片元件区域，再使用 CNN 进行分类。Kim YG^[21]提出一种后融合双流卷积神经网络，针对焊料区域检测，有效地提高了分类精度。这些方法在判定准则上得到了一定的优化，但并未解决一些前置工作的问题。目前，在工业视觉缺陷检测领域所应用的端到端的深度学习方法主要有两种，一种是利用正常样本数据训练的无监督学习方法，另一种是依赖于标签的监督学习方法。前者又可称为异常检测，目前已有的基于深度学习的异常检测方法主要有四个类别：基于距离度量、基于分类面构建、基于图像重构和结合传统方法的异常检测方法。它们的核心都是以大量的正常数据作为训练集训练模型，判断输入数据与训练集中的数据差异性即可^[22]。然而，此类方法只是将缺陷数据当成异常，而无法对具体类型做出区分，由于 SMT 产线中需要对检测到的缺陷结果进行统计，并且对缺陷类型做出了明确的分类，因此并不适用；而后者适用于缺陷类型明确的，在大量的正负样本数据训练后，通常能达到较好的效果。如与贴片元件缺陷检测类似的 PCB 裸板的缺陷检测任务中，Liu 等^[23]在 YOLO 算法的基础上提出了一种快速缺陷检测网络 MobileNet-Yolo-Fast，Ding 等^[24]结合 faster R-CNN 和特征金字塔提出了一种微小缺陷检测网络 (TDD-net) 来应对 PCB 缺陷目标小、难检测的难题。这些基于基于一阶段或者二阶段的经典目标检测框架开发的算法，同样也可以用于贴片元件的缺陷检测，然而目前已有的算法都无法满足 SMT 场景中对精度和速度的要求，因此迫切地需要一种更先进的算法模型。

1.3 研究内容

本文面向 SMT 贴片元件缺陷检测应用场景，针对当前采用的 AOI 质检技术进行分析，首先介绍了 SMT 质检的研究背景、研究现状以及相关的技术知识，然后紧抓以下两个关键技术上的痛点进行研究：一是 PCBA 局部图像到完整图像的拼接算法研究，二是贴片元件缺陷检测算法研究。具体研究内容如下：

1.3.1 基于 SURF 主要距离的主板图像拼接算法研究

AOI 设备在采集 PCB 图像时，由于相机视窗和对图像分辨率的双重要求，只能采取拍摄局部小图后再拼接成完整 PCB 图像的策略。在该工作中，首先对采图的路径进行了规划，通过将承载 PCB 板的置物台沿着 S 形的路线逐步地移动，每次拍摄当前处于视窗中的局部小图，以此逐步获取到所有区域的图片。由于需要对采集到的原图像和检测完成后的有标注图像都进行全局可视化，因此需要将所有的局部小图拼接成完整大图。

图像拼接的一般方法包括硬拼接和软拼接，前者通过直接测量图像间的重叠距离将像素简单叠加，后者通过基于特征匹配的拼接算法实现拼接。由于 AOI 系统拍摄局部图像时相机是按照固定的轨迹移动，且拍摄的环境较为稳定，使得相邻的图片之间存在着有规律的重叠。但 PCB 在拍摄过程中不可避免的会轻微震动，导致每次的重叠量发生变化，并且采集到图像的分辨率也很高。因此，直接使用硬拼接或者软拼接都存在着难以克服的弊端。本文首先针对软拼接中 SURF 算法流程进行分析，提出基于划分区域的特征点提取和基于主要范围的匹配对细筛选两种方案对原始算法进行改进，使得算法的速度和准确度都得到了较好的提升。然后从待拼接 PCB 图像的特点出发，充分结合硬拼接和软拼接，提出一种基于 SURF 主要距离的主板图像拼接算法，实验结果表明本文方法既拥有较快的拼接速度，又能够保证拼接成像的质量，能够很好地满足工业场景中的需求。

1.3.2 基于自适应协同特征的 DETR 缺陷检测算法研究

在该工作中，对 SMT 贴片元件印刷时产生的多种缺陷情况进行了分析，提出一种基于自适应协同特征的 DETR 缺陷检测算法研究。具体又分为两个阶段，第一阶段是缺陷数据集的建立，第二阶段是算法模型的设计。

在第一阶段中，由于工业场景中产品的良率较高，因此产生的可标注的主板缺陷数据非常稀少。另外经过实际调研发现，目前工厂中采用的 AOI 技术在检测出缺陷后，只会保存有问题的单一元件的图片。对该图片进一步处理，通过算法提取其中的感兴趣区域，去除冗余的背景后，得到了五类单一元件的缺陷图，包括偏移、缺件、立碑、侧立和少锡，每种缺陷的数量在 1000 张左右。在现有

情况下,采用了多种方法制作带标注的缺陷数据集。一是在少量的真实样本上进行人工标注;二是提出了一种 PasteChips 方法,该方法又包含两种内容,第一是将单一元件缺陷图粘贴到拍摄的 PCBA 局部图像中,并覆盖掉图中原先存在的无缺陷同类物体,第二是制作少量的主板模板图与单一元件缺陷图进行自动融合,可以用有限的数据生成任意数量的标注图片。通过以上方法,提出用于贴片元件缺陷检测的数据集。

在第二阶段中,考虑到现有检测算法的不足,包括传统的模板匹配算法和现有的基于神经网络的 YOLO 系列算法都无法很好满足工业场景中的需求,因此算法采用了基于 Detection Transformer (DETR) 的模型架构,实现了端到端的检测流程,并对模型做进一步的轻量化处理,在保障检测精度的前提下速度更快。具体来说,在高效实时目标检测器 RT-DETR 的基础上,针对 SMT 场景中的特别需求,提出了一种全新的自适应协同模块 ACBlock 替代部分传统的卷积模块,能够更好地关注到元件缺陷的形变特征;将 MPDIoU 和 InnerIoU 相结合设计了全新的 IMIoU 损失函数,它同时具备两者的优点,不仅具有相比于常见损失函数更快的收敛速度,且能够有效地检测小目标。综合以上方法,改进后的模型在自制的数据集上性能显著提高,并且优于其他所有的实时检测算法。同时,本方法也可以扩展到其他类似的领域,在一些公开数据集上也取得了不错的效果。

1.4 本文的组织结构

本文的具体章节安排如下:

第一章:绪论。介绍了 SMT 流程中 AOI 质检的研究背景,阐明了本文的研究意义,简要概述了当前 AOI 技术的研究现状,并紧抓现有技术的痛点,引出本文的研究内容。

第二章:相关技术。对 SMT 质检中涉及的相关技术进行了介绍。由于本文所研究的算法来自于实际的工程问题,并且算法的实施需要依赖于软硬件平台作为支撑。在本章中,首先介绍了算法平台,即 AOI 系统的结构和工作原理,并通过分析工作流程中所需要的关键技术,引出本文研究的两项内容,即图像拼接和缺陷检测算法,并且对这两项技术在实际中的应用进行了详细分析。

第三章:基于 SURF 主要距离的主板图像拼接算法研究。首先介绍了本文所使用数据的获取途径和预处理过程,然后说明了 PCB 图像整体的拼接流程,接着对硬拼接和基于 SURF 改进的软拼接进行深入分析,在此基础上,提出一种基于 SURF 主要距离的主板图像拼接算法,经过实验对比,验证了方法的有效性。

第四章:基于自适应协同特征的 DETR 缺陷检测算法研究。首先介绍了所提出算法的整体结构,包括基准模型 RT-DETR 的架构,以及本文所提出改进模

块的具体解释,包括自适应协同模块 (Adaptive Collaboration Block, ACBlock) 和 IMIoU (Inner and MPD IoU) 损失。然后在三个数据集,一是本文自行收集的 SMT 贴片元件缺陷检测数据集 (SMTC),二是公开数据集 DeepPCB 和 VisDrone2019,进行了大量实验,将结果与基准模型以及其他实时检测模型进行了对比和分析。

第五章:总结与展望。对本文所做的研究工作进行了总结,同时指出当前工作的不足与可改进之处,并展望了相关研究在未来的发展方向。

第2章 相关技术

本章主要介绍 AOI 系统工作过程中所设计的两项关键技术，包括基于 SURF 主要距离的主板图像拼接算法和基于自适应协同特征的 DETR 缺陷检测算法。由于在实际应用中，算法的实施都必须依赖集软件与硬件一体的 AOI 系统平台，故本章首先介绍了本文方法所依赖的 AOI 系统的结构和工作原理，然后紧抓工作流程中的技术痛点，先后对图像拼接和缺陷检测两项关键技术进行了分析，为下文提出新的解决方案做铺垫。

2.1 AOI 系统

本研究是在课题组与某知名消费电子制造企业深度合作的前提下开展的，为应对该企业 SMT 产线上对笔记本主板在回流焊环节后的贴片元件缺陷检测存在的问题，包括检测效率低、投入成本大等，对现有 AOI 系统进行多方位技术升级，设计出一款全新的 AOI 系统，以满足低成本、效率高、操作更方便的需求。需要说明的是，本文的工作重点在于改进 AOI 系统的两项关键技术，包括图像拼接和缺陷检测，其他的工作则由课题组整体完成。

2.1.1 组成结构

AOI 系统是集软件与硬件一体的自动化检测平台，在 SMT 产线上承担着 PCBA 质量检测的重要角色，涵盖了 PCB 检测过程中物料输送、图像采集和缺陷检测的全流程。本文的 AOI 系统主要包含以下五个模块：(1) 运动控制模块；(2) 光学成像模块；(3) 图像处理模块；(4) 缺陷检测模块；(5) 工控机。图 2.1 展示了 AOI 系统组成，图 2.2 是该系统的实物图。下面将进行详细介绍。

工控机是整个 AOI 系统的中枢，通过软件交互的方式调度其他各个模块的运行。考虑到在系统运行过程中需要涉及到大量的计算，对工控机的性能有一定的要求，因此，综合所有因素，本文按照表 2.1 中的指标对工控机进行配置。

表 2.1 工控机配置参数

指标	参数信息
处理器	Intel(R) Core(TM) i7-10875H CPU @ 2.30GHz
内存	16GB
硬盘	1TB
显卡	NVIDIA GeForce RTX 2060
显存	6GB

在运动控制模块，经过回流焊后的 PCBA 主板被传送轨道运送到 AOI 系统

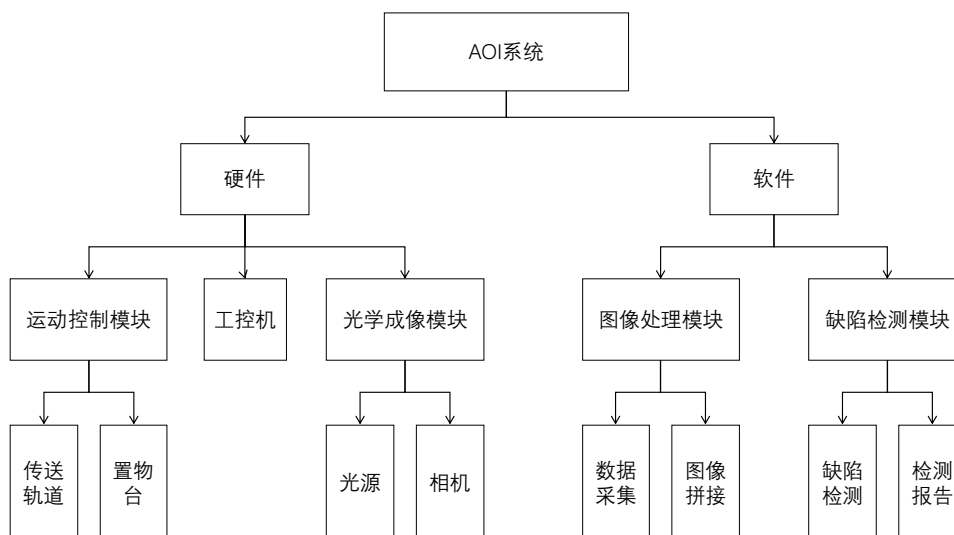


图 2.1 AOI 系统结构图



图 2.2 AOI 系统实物图

的置物台上，并且自动固定，以防在拍摄过程中发生抖动。同时，在图像采集环节中，由于相机视场和对图像分辨率的要求，承载着 PCBA 的置物台会在按照一定的路径运动，使得相机能够通过逐步拍摄局部图像从而扫描到主板的全部区域。本文对采图的路径进行了规划，制定了将置物台沿着 S 形轨迹运动的方式。由于不同型号的 PCB 主板存在差异，为了满足所有型号主板都能被完整地拍摄到所有区域，采用 9 行 7 列的采图方式。如图 2.3 所示，置物台的初始位置为点 $O(0,0)$ ，对应着图像的编号也为 $(0,0)$ ，首先向下方移动，在七次偏移后完成一列的拍摄，开始向右方移动一次，到达第二列的末端，接着方向转上，依此类推，经过 S 形的运动后到达终点 $P(8,6)$ ，完成 9 行 7 列共 63 张图像的拍摄过程。

在光学成像模块，光源和相机决定着拍摄图像的质量，也影响着后续环节中拼接和检测的效果。因此，对两者的选型都极为重要。对于光源而言，AOI 系统的要求是能够清晰稳定地拍摄到主板上每个元件的特征。其组成又分为光源控制器和 LED 光源^[25]，我们选取广东奥普特科技股份有限公司生产的型号为 KDY48W-E4X (X=4, 8, 16) 的光源控制器，配备“红色，绿色，蓝色”比例为

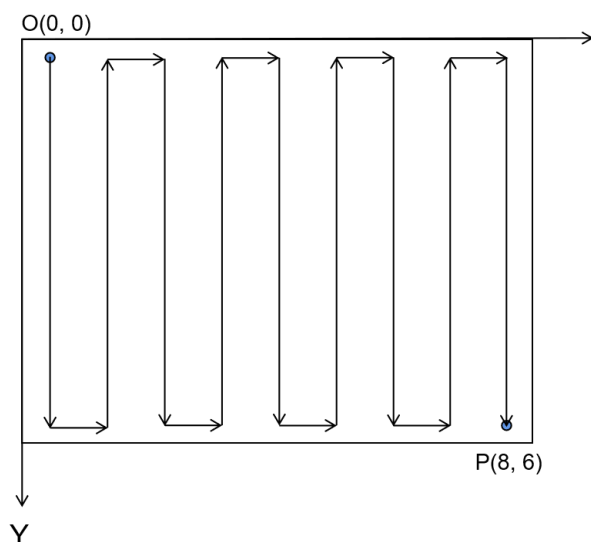


图 2.3 S 形采图路线

1:1:1 的 RGB 光源，实物图如图 2.4 所示，表 2.2 展示了具体参数。对于相机而言，采用高清的海康系列工业相机进行拍摄。

表 2.2 AOI 光源主要参数

光源指标	参数信息
光源型号	OPT-RIA298-RGB-K1
控制器型号	KDY48W-E4X (X=4,8,16)
开孔内径	100mm
开孔外径	298mm
功率	48w

在图像处理模块，负责对 PCBA 图像的数据采集以及之后的拼接工作。在前面提到，相机每次只能拍摄 PCB 的局部图像，工控机控制相机的运动并将拍摄的图像实时地存入文件系统的指定路径，以便下一步处理。当需要对图像进行拼接时，即可启动图像拼接流程，得到完整的 PCB 主板图像。

在缺陷检测模块，调用检测算法实时地对每一张采集到的图像进行缺陷检测，并且将检测的结果在屏幕上可视化，同时统计检测的结果，以便最终生成整板的缺陷检测报告。

2.1.2 工作原理

完成对 AOI 系统的设计工作之后，需要确定整个系统的工作流程。在系统应用之前。需要先对各个模块的参数进行调试，特别是光学成像模块，相机根据光源和拍摄距离确定光圈和焦距，以确保能采集到亮度适中且清晰的照片。应用时的流程图如图 2.5 所示，具体来说：首先系统启动，相机、光源和置物台复位到初始位置，传送轨道将待检测的 PCBA 运输到置物台上，完成放板。然后置物台按照预先设置的路径移动，每移动一次，相机拍摄一次当前图像并进行编号存



图 2.4 RGB 光源

人工控机文件系统。接着启动缺陷检测算法，对当前的局部图进行检测，并实时地将局部检测结果可视化。当完成对整板图像的采集后，缺陷检测也已经完成，此时根据检测结果来判断：如果板子有缺陷，就将统计的缺陷信息统计报告，并通过图像拼接得到完整图像，整板可视化后可以定位到缺陷位置，就判定为合格板，检测完成。

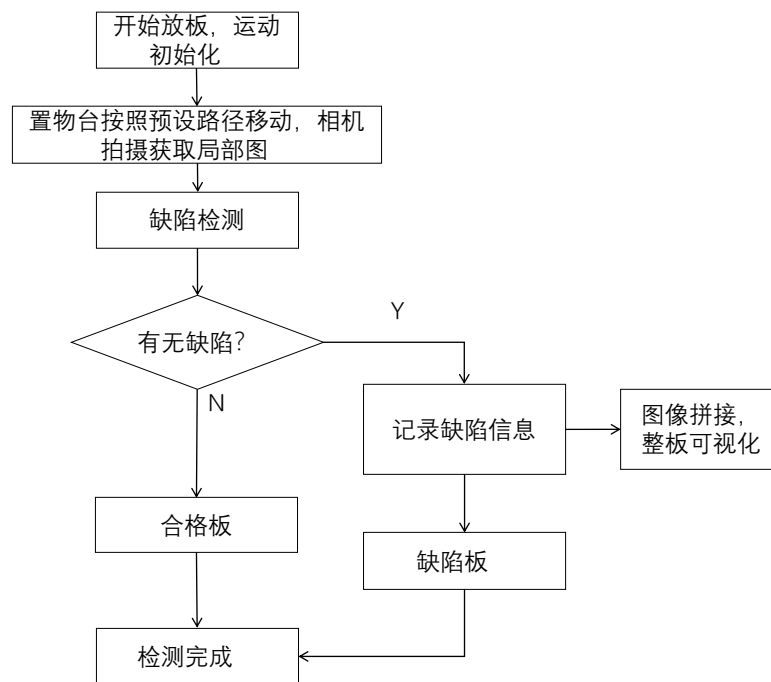


图 2.5 AOI 系统工作流程图

2.1.3 关键技术分析

以上两个小节介绍了本文 AOI 系统的组成结构和工作原理，可见 AOI 系统同时依赖于硬件和软件。受到成本问题的限制，不可能通过一味地提升硬件质量来升级系统，而那也属于是仪器科学与工程范畴。同样的，软件系统的设计也固然重要，但是其主要目的是提供足够的功能以提升人机交互时操作的简便性和体验感。本文的研究重点是，在基本配置达到要求的前提下，如何提升检测的效率。通过 AOI 系统的工作流程可以看出，有两种算法在发挥着重要作用，它们的速度和精准度都严重影响着整个 AOI 系统的效率：一是图像拼接算法，在整板可视化后可以给人以最直观的感受，并且可以据此快速定位到 PCBA 上有缺陷的位置，以便后续的返修操作；二是缺陷检测算法，关系着缺陷元件的检出成功与否，一旦判断失误，导致不合格的 PCBA 流入到后续环节中，将对产线资源造成不可挽回的损失。因此，本文针对这两项关键技术进行深入研究，接下来将介绍两者的相关技术的理论基础，为提出新的方案铺垫。

2.2 图像拼接理论基础

本节将介绍图像拼接的相关理论，其目的在于将 AOI 系统中采集的多张局部图像拼接成整板图像。主流的方法是基于特征点的拼接算法，本节重点介绍经典的 SURF 算法。

2.2.1 图像滤波

原始图像往往存在着不同程度的嘈杂噪声，这不利于图像拼接，因此需要先进行图像预处理工作。特别地，在 AOI 系统的工作环境中，所处理的 PCB 图像数据非常容易受到光照、温度变化等周围环境的影响，导致采集到的图像质量不一。图像滤波就是通过模糊处理的方法，使图像变得平滑，这不仅减少了显著噪声点，也降低了后续 SURF 算法提取特征点过程中的计算量。图像滤波又分为空域滤波和频域滤波两种类型，一般来说，频域滤波的算法较为复杂，且需要经过大量的运算，计算时间长。因此，针对 PCB 图像的滤波处理，本文全部采用空域滤波，其本质是使用滤波模板对目标像素点及其邻域内像素值进行再计算，主要有以下几种类型：均值滤波、中值滤波、高斯滤波和双边滤波。

1. 均值滤波

均值滤波的思想是将当前像素点 $P(x, y)$ 的灰度值置为邻域内所有像素值的均值，邻域大小为模板的尺寸大小 $(N \times M)$ ，得到新像素点 Q 的运算过程可表示为：

$$Q(x, y) = \frac{1}{N \times M} \sum_{f \in S} f(x, y) \quad (2.1)$$

均值滤波将某一点的噪声干扰平均分担到该点的整个邻域，使图像变得平滑，且随着模板变大，图像会变得更加模糊。该算法实现较为简单，但会破坏图像的细节部分，特别是轮廓边缘部分受到明显损失^[26]。

2. 中值滤波

中值滤波针对图像中的每个像素点 $P(x, y)$ ，首先根据所选的模板对该点邻域内的所有像素值进行排序，然后用序列的中间数值来替代原始像素值，用公式表示为：

$$g(x, y) = \text{Median}f(x, y), (x, y) \in S \quad (2.2)$$

其中， $g(x, y)$ 为 P 点经过滤波后的像素值， S 为邻域，通常邻域大小为奇数，如 3×3 、 5×5 等。中值滤波的计算也比较简单，但不同于均值滤波，它对能较好地处理孤立噪声点，如椒盐噪声，并且对边缘的破坏较轻^[27]。

3. 高斯滤波

一般来说，对于图像中某一个像素点 $P(x, y)$ ，其周围的点与其距离越近，联系应该更密切，反之则越疏远。基于这种思想，高斯滤波对邻域内不同的像素点根据距离的不同施加重权，且权值符合正态分布^[28]。根据图像的性质，选择二维高斯函数作为模板：

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.3)$$

其中， σ 为标准差。通过高斯函数可以生成高斯滤波模板，而模板的各点值与 σ 有着重要关系，直接影响到滤波的效果。根据高斯函数的分布情况， σ 越小，表明近点的权重越大，远点的权重越小； σ 越大，表面近点的权重越小，远点的权重越大。由于近点的值始终大于远点，就导致 σ 越大，函数值的分布曲面就越平缓，对图像的平滑效果越明显，趋近于均值滤波。以 3×3 的模板为例， σ 为 0.8，假设中心点坐标为 $(0, 0)$ ，则生成的高斯模板如图 2.6 所示。

(-1, 1)	(0, 1)	(1, 1)	➔	1	2.1842	1
(-1, 0)	(0, 0)	(1, 0)		2.1842	4.7707	2.1842
(-1, -1)	(0, -1)	(1, -1)		1	2.1842	1

图 2.6 生成高斯模板

经过归一化取整后得到整数型的高斯模板见图 2.7。

$\frac{1}{16}$	1	2	1
	2	4	2
	1	2	1

图 2.7 整数型高斯模板

可以观察到，高斯模板上距离中心点不同距离的位置权值不同，将其作为滑动窗对图像上的每个像素点做卷积运算，得到的结果替换原本的像素值，就完成了高斯滤波的过程。高斯滤波可以很好地处理高斯噪声，但对边缘信息保存效果较差。

4. 双边滤波

同高斯滤波一样，双边滤波也采用了加权的思想。然而，单一的距离加权更适合像素值变化平缓的区域，在具有明显变化的区域，例如图像中的边缘，中心点与周围部分像素联系不大，直接通过距离加权就破坏了这类信息。因此，双边滤波又考虑到了像素值差异的影响，增加了像素值权重项^[29]。用 G_s 表示距离权重， G_r 表示像素值权重，两者分别为空域和像素值域的二维高斯函数，双边滤波可表示为：

$$g(x, y) = \frac{\sum_{k,l} f(k, l) \omega(x, y, k, l)}{\sum_{k,l} \omega(x, y, k, l)} \quad (2.4)$$

其中， $g(x,y)$ 为滤波后的输出像素点， $f(k,l)$ 代表输入像素点， $\omega(x,y,k,l)$ 为空间域函数与像素值域函数的乘积，公式表示为：

$$\omega(x, y, k, l) = \exp \left[-\frac{(x-k)^2 + (y-l)^2}{2\sigma_d^2} - \frac{\|f(x, y) - f(k, l)\|^2}{2\sigma_r^2} \right] \quad (2.5)$$

其中， σ_d 是空域高斯函数 G_s 的标准差， σ_r 为值域高斯函数 G_r 的标准差， $f(x,y)$ 为函数作用的中心点， $f(k,l)$ 为邻域内的任一点。双边滤波对高斯噪声的抑制效果较好，并且相对于高斯滤波，对于边缘信息的保护效果更好。

2.2.2 SURF 特征检测算法

SURF 算法用于检测图像中的特征点，并且对于特征点的尺度变换、旋转变换不敏感，通过将两幅图像中的相似特征点配对，以达到图像配准的目的。SURF 算法的一般流程为：构建尺度空间、特征点精确定位、确定特征点主方向和构建特征描述子。

1. 构建尺度空间

为了提取图像中的鲁棒性特征，首先要在原图的基础上建立多尺度空间。SIFT 直接将灰度图的像素值与高斯差分函数 (Difference of Gaussian, DoG) 进行卷积运算，得到 DoG 尺度空间。而高斯核函数是一种二阶函数，导致出现大量的数学运算，消耗了大量时间。为了简化运算，SURF 引入积分图像^[30]的概念，如图 2.8 所示。

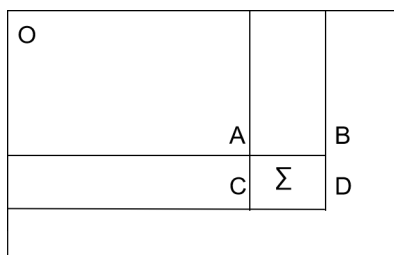


图 2.8 积分图像

对于图像中某一点 $P(x, y)$, 其在积分图像上的值为 $I_{\Sigma}(P) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i, j)$, 表示由原点和该点组成的矩形区域内所有像素的总和。然后，对于上图中的任意矩形区域，其内的像素和为：

$$\Sigma = A - B - C + D \quad (2.6)$$

其中，A, B, C, D 分别表示该点的在积分图像中的值。计算像素和的过程不再与区域大小相关，而只需要三次加减运算，极大地提升了运算效率。

在此基础上，采用 Hessian 矩阵建立盒式滤波器 (Box Filter)，来近似替代高斯二阶滤波器，其中 Hessian 矩阵的数学表达式为：

$$\mathbf{H}_{f(x,y)} = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} \quad (2.7)$$

该行列式的值为：

$$\det(\mathbf{H}) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad (2.8)$$

如果将 $\det(\mathbf{H})$ 作为判别项，则可以根据其正负性得出点 (x, y) 是否是极值点。图像在高斯滤波后，构造海森矩阵：

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (2.9)$$

其中， L_{xx} 为图像上一点 $I(x, y)$ 与二阶高斯模板的卷积，其公式为：

$$L_{xx}(x, y, \sigma) = \frac{\partial^2 G(x, y, \sigma)}{\partial x^2} * I(x, y) \quad (2.10)$$

L_{xy} 和 L_{yy} 的计算过程与之相仿。然而上述过程在实际计算过程中，并不会真的采用高斯二阶模板对图像卷积，而是采用盒式滤波器近似海森矩阵中的高斯滤波模板，图 2.9 展示了两者的近似过程：

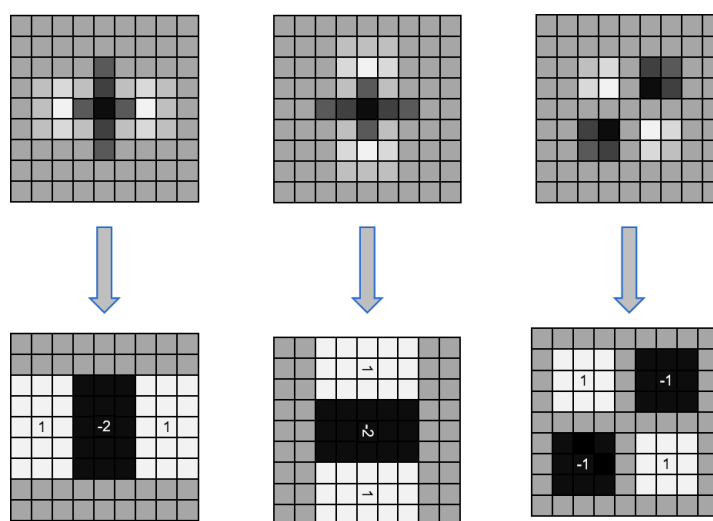


图 2.9 盒式滤波器

图中 (a)、(b)、(c) 三图分别用盒式滤波器的三个滤波模板 D_{xy} 、 D_{xy} 和 D_{xy} 近似代替相应的高斯二阶模板 L_{xy} 、 L_{xy} 和 L_{xy} 。图中不同颜色的块代表不同的权重，黑色权重为-2，白色权重为 1，灰色权重为 0。从几何意义上说，可以看成用平面去代替曲面，从而极大地提高计算效率。接着，用 Hessian 矩阵的行列式来计算：

$$\det(H) = D_{xx}D_{yy} - D_{xy}^2 \quad (2.11)$$

为了平衡近似代替的误差，通常需要以 0.9 的加权系数来缩放 D_{xy} 。

滤波器的作用在于对图像进行卷积以提取图像特征，通过改变盒式滤波模板的大小，并使用不同尺寸的模板对图像进行模糊处理，可以生成不同程度的模糊图像，从而建立图像的尺度空间。特别地，传统的方式对图像用串行的方式多次滤波，在上采样时上层图像依赖于下层图像，导致效率低下，而用不同尺寸的模板可以并行地进行滤波处理，大大节省了时间，滤波过程如图 2.10 所示。

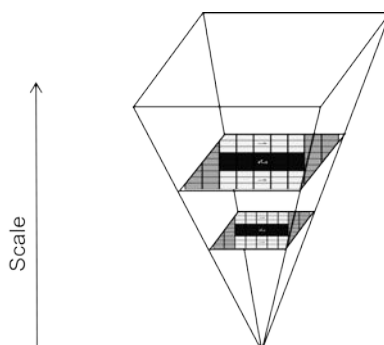


图 2.10 盒式滤波

2. 特征点精确定位

通过计算得到尺度空间内某一点的近似海森矩阵行列式的正负性，可以判断该像素点是否为极值点。然后再将其与 $3 \times 3 \times 3$ 的邻域内其他 26 个点比较，如果为最大值或者最小值，就可以初步判断为关键点。由于像素值是离散的，通过滤波模板拟合的连续空间内的极值点可能处在两点中间，因此需要采用二次插值来确定亚像素级特征点的精确位置。

3. 确定特征点主方向

在特征点的半径为 $6S$ (S 为特征点所在的尺度值) 的圆形邻域内，统计所有其他点在垂直和水平两个方向的 Harr 小波 (Harr 小波模板边长为 $4S$) 响应，再对响应值进行高斯函数加权，得到该点在两个方向上的方向权重。将整个圆域等分为 6 份扇形区域，分别统计通得到该区域内的最大分量，以最大分量的方向为特征点的主方向。如图 2.11 所示。

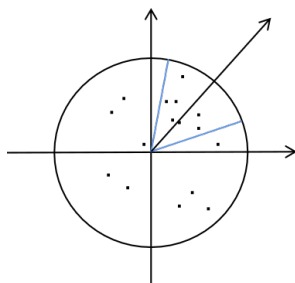


图 2.11 特征点主方向

4. 构建特征描述子

如图 2.12 所示，在以特征点为中心，周围 $20S \times 20S$ 的正方形邻域内建立笛卡尔坐标系，Y 轴为特征点主方向。再将该区域划分为 4×4 个子块，每个子块大小为 $5S \times 5S$ ，统计每个子块内像素点在垂直和水平方向的 Harr 小波响应值之和，包括四个分量 $\sum d_x$ 、 $\sum d_y$ 、 $\sum |d_x|$ 和 $\sum |d_y|$ ，得到一个 $4 \times 4 \times 4$ 共 64 维的特征点描述向量。最后对其归一化处理以消除光照影响，得到特征点描述子。

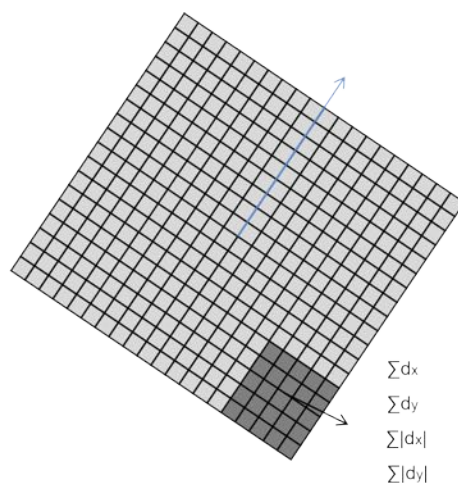


图 2.12 特征点描述子

2.3 缺陷检测理论基础

目标检测算法的目的在于检出图像中感兴趣的对象，并确定其位置和类别信息。本文研究的 SMT 贴片元件缺陷检测任务与之匹配。目前，基于深度学习的目标检测又分为基于卷积神经网络的目标检测算法和基于 transformer^[31]的目标检测算法。考虑到不同模型的特性，经过比较后选择基于 transformer 的 DETR 模型作为基准架构，以下将对相关的技术进行介绍。

2.3.1 基于卷积神经网络的目标检测算法

基于卷积神经网络的目标检测算法一般分为两阶段和单阶段两种类型，前者的思想是将检测分为两步，即先定位，再识别，通常具有较高的检测精度，但是速度较慢，不适合实时应用，代表作有 R-CNN^[32]，Fast R-CNN^[33]和 Faster R-CNN^[34]等；后者只用一个网络就解决了定位和识别两个任务，在实时性上大幅提升，代表作有 YOLO^[35]，SSD^[36]等。以下主要介绍经典的 R-CNN 系列和 YOLO 系列算法。

1. R-CNN 系列算法

R-CNN 算法于 2014 年被提出，开创了 CNN 作为基准架构的目标检测系列算法的先河，极大地提升了目标检测的效率。作为典型的两阶段算法，R-CNN 对输入图像处理的主要流程如下：

- (1) 生成大量候选区域（约 2000 个）；
- (2) 使用 CNN 提取候选区域的特征；
- (3) 使用分类器识别提取的特征；
- (4) 使用非极大值抑制（Non-Maximum Suppression, NMS）去除冗余的候选框；
- (5) 输出检测到的目标，包括类别和位置信息。

图 2.13 是 R-CNN 算法的图解过程。然而，R-CNN 对每个候选区域都要做一次特征提取与分类，相当于将一张图像分成了非常多的子图，且对于不同的类别需要建立不同的分类器，这导致了极大的计算量和内存消耗。

2015 年提出的 Fast R-CNN 算法，是对 R-CNN 算法的改进。它将 R-CNN 中三个独立的任务（候选区域生成、特征提取和分类）整合到一个网络中，大大提升了算法的效率。具体来说，直接将整张图像输入到 CNN 中，在输出的特征图上截取候选区域，再经过 ROI 池化层调整到同一尺寸，然后经过一系列的全连接层直接得到物体的类别和位置信息。不过在特征图中截取候选区域的步骤依旧复杂，Fast R-CNN 对每张图像的检测时间仍然需要约 2s 之久。

Faster R-CNN 对算法进行了更为高效的优化，提出 RPN（Region Proposal Networks）生成先验框以替代传统的候选区域生成方法，使得 Fast R-CNN 中最

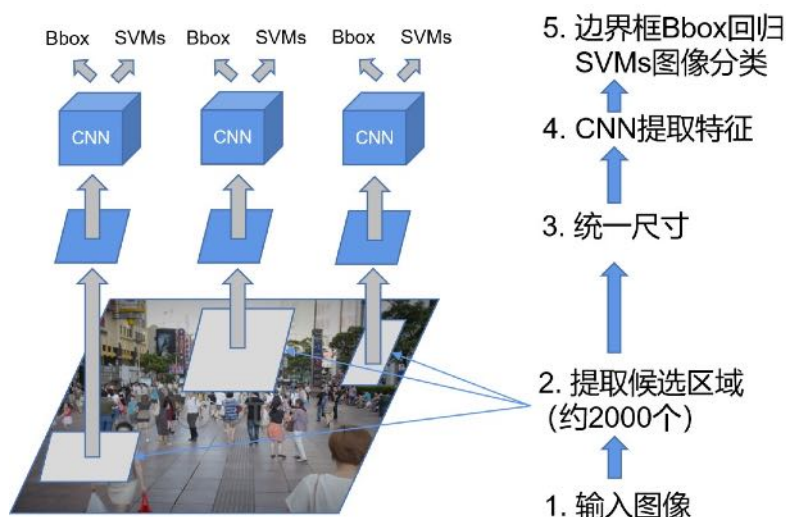


图 2.13 R-CNN 算法流程图

耗时的操作得到了极大提升。

2. YOLO 系列算法

YOLO 算法是典型的一阶段目标检测算法，从 2015 年 Joseph Redmon 提出 YOLO 最早的版本至今，已经迭代了多个版本，目前最先进的 YOLO 算法已经不单在速度上具有优势，在检测精度上也达到了高水平。

YOLOv1^[35]首次将目标检测当做回归任务来处理，它将输入图像划分为 $S \times S$ 个网格，并为每个网格单元格分配一个负责预测该单元格中目标的责任。YOLOv1 算法使用卷积神经网络 (CNN) 提取图像特征，每个网格单元格预测多个边界框 (bounding box)，每个边界框包含目标的类别、位置 (x,y,w,h) 和置信度信息。最后通过非极大值抑制算法去除冗余的边界框，得到唯一的预测结果，如图 2.14 所示。然而，当检测密集型物体或微小物体时，由于网格大小固定，同一个网格可能存在多个目标，效果不佳。YOLOv2^[37]从检测精度、速度和识别对象数量三个方面升级 YOLOv1，采用 Darknet-19 作为主干网络，并且提出一种训练联合分类和检测的方法，将可检测目标扩展到 9000 个。YOLOv3^[38]引入了残差结构，将主干网络的深度拓展到 53 层得到 Darknet-53，并且引入 FPN 融合多尺度特征信息，模型的整体性能得到提升，并且小目标的检测能力显著增强。在此之后，YOLO 的基本架构已经确立，后续的升级版本都是在模型训练、数据优化等方面提出不同的 tricks 以提升模型的性能。

YOLOv4^[39]使用了较多的调优手段，包括加权残差连接、马赛克数据增强、余弦退火策略、Mish 激活函数、CIoU 损失函数等，在主干网络和特征金字塔部分都做了优化，使得 YOLOv4 的检测精度和速度达到了较好的平衡，同时也降低了训练时的硬件要求，但整体参数量较大。YOLOv5 诞生于 YOLOv4 提出后不久，它不再基于以往 YOLO 算法中的 Darknet 开发，而是由 Ultralytics 团队移植

到了 Pytorch 平台，并且提出了一些新的改进思路，性能得到了进一步提升，且部署更为简便。YOLOv6^[40]由美团视觉人工智能部提出，在网络结构、训练策略等方面做出了多项改进，如使用了更先进的主干网络 EfficientRep，对于工业应用非常友好。YOLOv7^[41]的改进点主要在于模块重参数化和动态标签分配策略，这使得训练时间增长，但推理时的速度超越了以往的算法。YOLOv8 是 YOLO 系列算法最先进的版本，它仍然出自 YOLOv5 的作者 Ultralytics 团队，它融合了目标检测发展至今的多项 SOTA 技术，具体创新包括一个新的骨干网络、一个新的 Ancher-Free 检测头和一个新的损失函数，使得模型在更加的轻量化的同时，检测性能继续提升，工业应用的前景巨大。

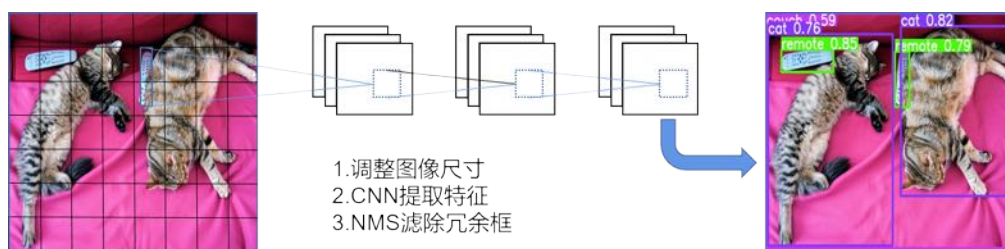


图 2.14 YOLO 算法流程图

2.3.2 基于 transformer 的目标检测算法

尽管基于 CNN 的目标检测算法已经取得了不错的效果，但是在以往的架构中，仍然采用的是间接预测的方式，如 RCNN 系列使用了候选区域，YOLO 系列大多采用了锚框，这些算法无一例外地会生成大量的冗余预测框，需要通过 NMS 算法进行后处理才能得到最终的结果。然而，NMS 本身就非常耗时，使得这些模型的优化变得非常局限。根本原因在于 CNN 每次只能对图像进行局部处理，而无法获得当前视野范围外的信息，因此，只能基于多个局部的视野生成若干范围内的预测。Transformer 的一个突出能力在于全局建模，在自然语言处理领域，文本信息以序列的方式直接输入到 Transformer 的编码器中，而后通过注意力机制建立全部位置的联系。这种方法也同样可以运用到计算机视觉领域，ViT(Vision Transformer)^[42]开创了先河，它将图像分成多个块，再将每个块投影成固定长度的向量序列作为 transformer 的输入，实现了基于 transformer 的图像分类。2020 年由 Meta AI 提出的 DETR(Detection Transformer)^[43]将 transformer 引入目标检测领域，将目标检测看做集合预测的问题，提出了一个十分简洁的目标检测框架，即 CNN 提取基础特征，送入 transformer 做关系建模，得到的输出通过二分图匹配算法^[44]与图片上的 ground truth 做匹配，直接输出物体的预测框，真正实现了端到端的目标检测，算法流程如图 2.14 所示。

然而，在借助 transformer 强大建模能力的同时，DETR 的也受制于 transformer 高计算量。究其原因，一是 transformer 在处理图像时，注意力机制关注了全部

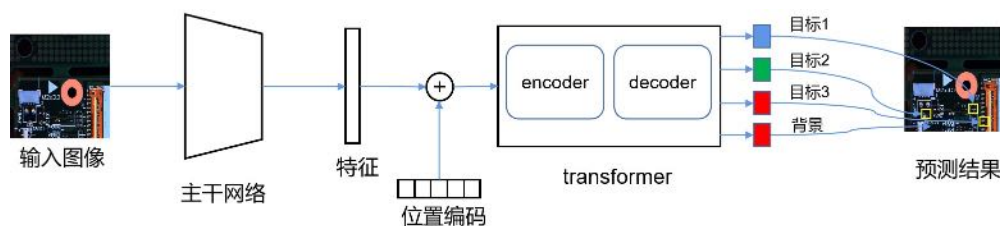


图 2.15 DETR 算法流程图

的像素点，一开始权重均匀分布在图片的每个像素上，需要大量的训练才能让权重集中在小区域上；二是编码器上的计算复杂度是像素数量的平方级，无法使用多尺度的特征图。Deformable DETR^[45]提出可变形注意力来改善 DETR 的不足，不同于传统注意力机制中将图像的每个像素点之间都做一次交互，可变形注意力通过采样的方式，让每个特征像素只需要与部分采样获得的像素交互即可，大大加速了模型收敛。同时，计算量减小之后多尺度特征融合也可以得到使用，这提升小目标检测的能力。为了进一步提升模型的检测性能，后续一系列的算法在 DETR 的基础上做出了改进，如 DAB-DETR^[46]对 DETR 中定义模糊的可学习查询做出了具体解释，将其分为位置查询和内容查询；DN-DETR^[47]加入对比去噪的训练方法以改善二分图匹配的不稳定性；DINO^[48]通过混合查询选择优化原本粗糙的查询。经过近几年的迅猛发展，DETR 系列算法的性能得到了极大地加强，在多个热门公开数据集（如 MS COCO、Object365 等）上的检测精度上一度超越了所有基于 CNN 的目标检测器。

第 3 章 基于 SURF 主要距离的主板图像拼接算法

3.1 引言

图像拼接算法在整个 AOI 系统中的作用是将拍摄的局部图像拼接成完整的 PCB 图像，而获取整板图像又有两个用处，一是与模板图像进行比对，标定出待检测的区域；二是缺陷检测完成后的整板可视化。基于第二章内容的分析，传统拼接算法存在着速度慢、成像质量低等多项不足，本文研究的拼接算法基于 AOI 系统平台所采集的 PCB 图像，对图像拼接的一般流程进行优化。另外，在本文特定的算法应用环境中，本文方法巧妙地结合了采集图像的本身的特征，以较低的成本实现了 PCB 图像拼接的高速度和高质量。

图像拼接又分为硬拼接和软拼接，硬拼接依赖既定的偏移量将两幅图像的像素直接叠加实现拼接，软拼接则依赖于拼接算法将两幅不同视角的图像置于统一的坐标系中，再通过像素映射实现拼接。本文的目的在于实现 AOI 系统采集到的 PCB 图像的拼接过程，通过实验发现，单纯的硬拼接和软拼接都无法满足要求，且存在着各自难以克服的弊端。因此，在本章中结合两种基本方式，提出一种基于 SURF 主要距离的主板图像拼接算法，并通过实验证明了方法的有效性，能够很好地满足实际需求。具体来说，主要对以下内容进行了研究：(1) 基于 SURF 匹配对的主要距离法，首先提出基于划分区域的特征点提取方法和基于主要范围的匹配对细筛选方法对 SURF 算法的进行改进，以相邻图像间的主要距离作为偏移量，弥补了硬拼接过程中偏离量无法实时调整的不足；(2) 基于空域的像素加权融合法，对光照不均造成的拼接缝两边明显色差进行有效处理。

本章结构安排如下，3.2 节对本文特定场景下待拼接图像的采集过程进行了说明。3.3 节对图像进行初步分析并且做预处理消除噪声。3.4 节介绍了本文采用的图像拼接方法，首先对整体流程进行介绍，然后分别分析了硬拼接和基于 SURF 改进的软拼接及其不足，并提出本文的拼接方法。3.5 节是实验结果与分析，证明了本文方法的有效性。3.6 节对本章内容做了简要总结。

3.2 图像采集

由于贴片元件的规格通常非常小，在 PCB 图像中属于微小目标，因此，要想对其进行缺陷检测，就必须确保拍摄的图像具有足够的清晰度。众所周知，光学设备的价格通常非常高昂，直接使用超高分辨率的相机拍摄整板图成本太大，因此，采用普通相机多次拍摄再拼接的方法是工业中最可行的。在主流的 AOI

设备中，采图的方式一般有两种：一是固定置物台，移动相机；而是固定相机，移动置物台。考虑到前者可能出现相机快速移动导致图像抖动的问题，本文 AOI 系统采用的是后者。由于移动范围是有限的，且相机只需扫描到置物台上的全部区域就不会导致拍摄不全。因此，数据采集的路线是可以事先调试好的。经过调试，设定采集路线：将整个置物台区域划分成 9 行 7 列共 63 个相同大小的子区域，且每个子区域在交界处都留有一定的重叠，按照 S 形的路线移动置物台逐步获取到所有图像。如图 2.3 所示，本文将在列中移动的方向定义为 Y 轴，将在行中移动的方向定义为 X 轴。

3.3 图像预处理

3.3.1 实采图像分析

根据本文 AOI 系统的设置，对每张 PCB 拍摄会采集到 63 张图像，在文件系统中的文件名标号依次为 0,1,2,3...63，按照采图过程可以分为 9 行 7 列，每行有 7 张图像，每列有 9 张图像。经过测量，每张图像的尺寸都为 3648x5472，占存为 29.8MB，因此每张 PCB 板的图像共占存约 1.83GB。进一步地，我们尝试测量相邻图像间的偏移距离。理论上来说，由于置物台每次在 X 轴和 Y 轴中移动的距离是标定好的，因此可以获取两种移动方式下相邻图像的重叠像素量，本文将其分别定义为行偏移和列偏移，如图 3.1 所示。然而，由于置物台移动时的轻微震动等原因，行偏移和列偏移不可能保持不变。经过测量，行偏移为 70 ± 20 像素距离，列偏移为 1880 ± 20 像素距离。

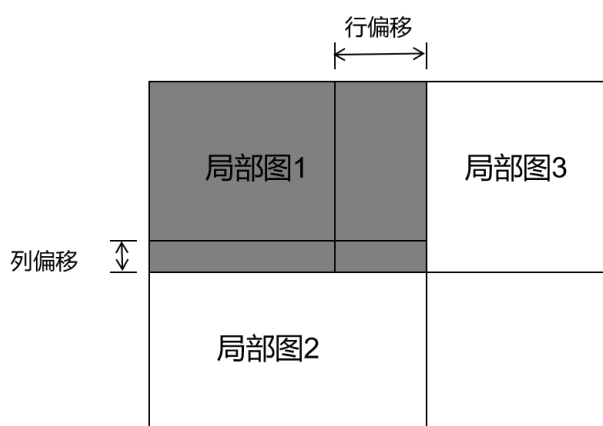


图 3.1 行偏移和列偏移

3.3.2 图像去噪

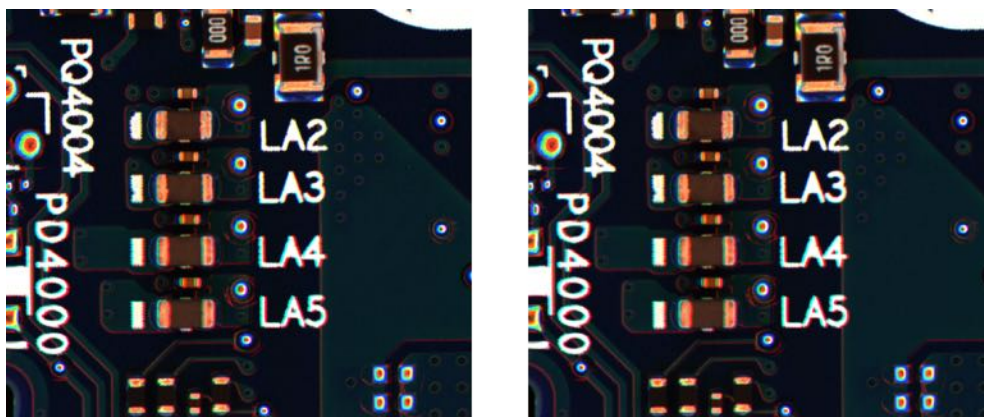
考虑到由于采集到的图像不仅作为本章拼接算法的对象，还是后续缺陷检测算法的输入数据。因此，对图像进行预处理是非常必要的，目的是去除由外部环境干扰造成的噪声。下面将对可能存在的噪声情况进行分析，以确定最佳的去

噪处理方案。

(1) 由于在图像采集过程中, PCB 直接暴露在现实的开放环境中, 外部光照强度的变化会造成图像亮度不均匀, 另外贴片元件的密集排列导致自身噪声相互影响, 都会引起高斯噪声。

(2) PCB 裸板本身在制造过程中容易出现喷漆不均匀而产生部分小漆点的情况, 造成局部的椒盐噪声。

单一的滤波器只对特定种类的噪声具有良好效果。为了处理 PCB 图像中可能存在的多种噪声, 本文提出了一种复合滤波解决方案, 该方案先后结合了双边滤波和中值滤波。先使用双边滤波, 在滤除高斯噪声的同时尽量保留图像的高频细节, 如贴片元件的轮廓信息、PCB 的标志部位等; 再利用中值滤波, 去除椒盐噪声, 这可能包括由环境因素影响的孤立坏点。经过以上处理后, 得到的效果图如图 3.2 所示, 注意, 为了能够观察到图像的视觉特征变化, 此处只截取了采集图像的一部分区域。



(a) 原始图像

图 3.2 图像预处理效果

3.4 图像拼接算法设计

图像拼接技术就是将数张有重叠部分的图像(可能是不同时间、不同视角或者不同传感器获得的)拼成一幅更大型的无缝高分辨率图像的技术^[49]。由于本文提出的拼接算法是应用在实际场景中的, 需要针对实际问题具体分析。在本节中, 首先确定了图像拼接的流程, 然后提出三种解决方案, 包括硬拼接、基于 SURF 改进的软拼接和基于 SURF 主要距离的图像拼接算法, 通过对比实验确定了最终方案。

3.4.1 图像拼接流程

根据上节内容可知,在本文的应用场景中,图像的采集方式是确定的。而图像拼接的对象是两张图像,因此无论采用哪种拼接算法,对最终整板图像的拼接流程都是相同的。基于图像采集的路线,对采集的 63 张图像采用先列后行的拼接方式拼接成完成 PCB 图像。同时,考虑到图像的在硬盘中的读和写都需要时间,且图像越大耗费时间越长,因此,为了节省时间,在拼接过程中采用了以下两种规则:一是边拍摄边拼接,只要满足有两张可拼接的图像就启动拼接算法,而不用等待文件全部写入硬盘;二是不保留拼接过程中的所有中间图。具体的拼接流程如图 3.3 所示。

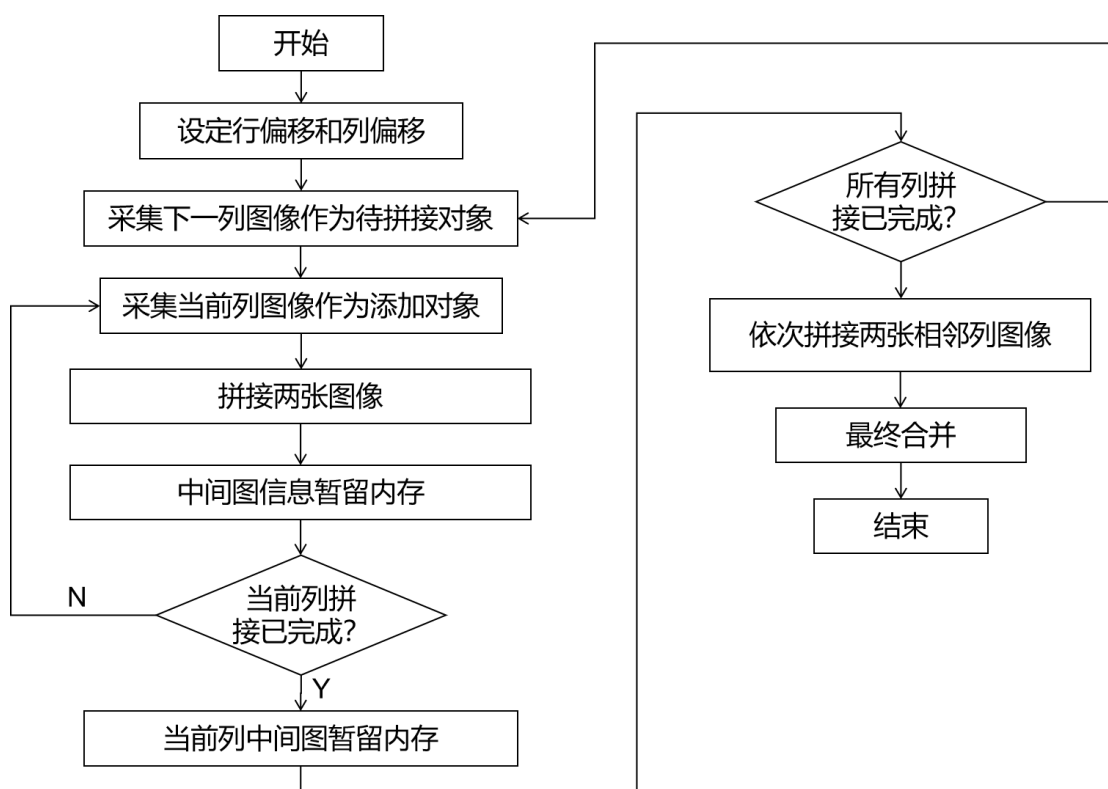


图 3.3 PCB 图像拼接流程

3.4.2 硬拼接

由上节所述,行偏移和列偏移是可以测量的,但是存在误差量。进一步地,每次拍摄的偏移误差也是在变化的,且暂时还无法做到无误差。因此,直接根据测量的像素距离将图像叠加,即硬拼接的方式,会导致较大的误差。在最坏情况下,由于误差的多次累积,行偏移会达到 160 像素距离,列偏移达到 120 像素距离,这对最终的拼接结果是不可接受的。在本节中,仍采用硬拼接的方式进行尝试,不同的是,我们对行偏移和列偏移的确定采用了多次测量取平均值的方法。表 3.1 显示了在 10 次拍摄过程中,对每次拍摄的第一行和第一列的相邻图像偏移距离,通过 PhotoShop 软件手工测量得到的结果,取均值后设定行偏移为

67Pixels, 列偏移为 1888Pixels。

表 3.1 行偏移和列偏移测量

拍摄轮次	行偏移 (Pixels)	列偏移 (Pixels)
1	65	1898
2	67	1888
3	64	1880
4	69	1890
5	73	1883
6	64	1884
7	62	1898
8	65	1894
9	65	1886
10	68	1888

硬拼接基于 OpenCV 库中的 concatenate 函数, 每次拼接两幅相邻图像。具体来说, 对于列拼接, 如图 3.1 中的局部图 1 和局部图 2, 按照设定的列偏移, 将 1 的重叠区域截断, 再将 2 与截断后的 1 通过 concatenate 函数拼接, 得到两张图拼接的中间图, 再将此中间图作为下一次拼接的局部图 1, 与下一张相邻的图像拼接, 依次类推, 完成整个列的拼接; 对于行拼接, 同理。采用平均偏移量的方法对图像拼接质量有一定程度的提升, 但仍然存在着较大的误差。

3.4.3 基于 SURF 改进的软拼接

考虑到硬拼接时采用的偏移量具有误差, 且误差大小无法控制, 使得硬拼接的不稳定性非常大。如果误差较大, 就会导致拼接完成后的图像出现明显的拼接裂缝, 这在高精度的场景要求中无法满足需求。本节采用 SURF 拼接算法用于相邻图像间的拼接, 分别提出一种基于划分区域的特征点提取方法以替代原始的特征提取方案和一种基于主要范围的匹配对细筛选方法以剔除错误的匹配对, 实现了基于 SURF 改进的软拼接算法。下面将以 AOI 系统实际采集的两张相邻 PCB 图像为例, 具体介绍 SURF 算法的应用, 流程如图 3.4 所示, 其中图像预处理阶段已在前文中做了介绍。



图 3.4 相邻图拼接流程

1. 基于划分区域的特征点提取

AOI 系统中的 PCB 图像在拍摄过程中是通过置物台的平移逐步获取的, 因此相邻图像间具有非常良好的平移特性, 同时移动过程中的轻微震动导致小角度的偏移, 又使得两幅图像间出现一定的旋转变换。PCBA 主板上的存在着各种

元器件、孔位、锡膏和字符等，经过平移旋转变换后能够产生大量稳定的点特征，利用好图像间重叠部分的这些特征，将对拼接算法有着非常重要的意义。针对点特征提取的算法主要有 SIFT^[10-11]、SURF^[12]、BRISK^[50]和 ORB^[51]。其中 SIFT 提取的有效特征点最多，但速度最慢，不利于图像的实时拼接；ORB 的速度最快，但是提取到的特征点较少，在高清图像拼接时稳定性不足；BRISK 在处理较大模糊图像时表现最好，也不适用于本文场景^[52]；SURF 在 PCB 图像中表现最佳，兼顾了速度和稳定性。

另外，由于 PCB 图像的采集过程具有较强的稳定性，且相邻图之间的重叠只发生在图像四周的一小部分区域，偏移误差也可以通过测量得到大致的范围。因此，为了充分提高拼接的效率，我们基于不拆分重叠区域的准则，在行拼接和列拼接的过程中，分别将图像裁剪为三个区域。以行拼接为例，如图 3.5 所示，当前图像最多只需要考虑同行中左右两张相邻图像，其重叠只与图像的长度有关，因此采用划分区域的方案，将原图像（每张图像的尺寸为 3648x5472）拆分为左中右三个区域，其中重叠范围是左右区域的子集。

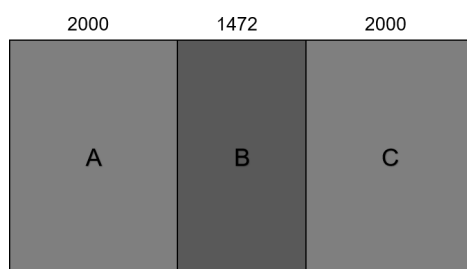


图 3.5 行拼接区域划分

根据偏移量的变化范围 (1880 ± 20)，行拼接时左右区域的长设为 2000，中间长为 1472 的区域面积不参与特征检测，而是直接保留原像素，在 SURF 拼接完成后直接添加到最终的图像中。

图像切分后，需要使用算法的区域面积变小很多，使得特征检测的效率提升，同时也抑制了无重叠区域中特征的干扰。图 3.6 展示了待检测区域的 SURF 特征点：

海森阈值是用于确定图像中的感兴趣区域和特征点的一个重要参数，决定了 SURF 算法在图像中检测何种大小的特征点或兴趣点，阈值越大，检测的特征点尺度越大。左图为阈值 2000 时的情况，一共检测到 3000 个特征点，分布过于密集，为了提升匹配的效率，需要对阈值进行重新设计。经过实验，发现阈值设为 10000 时效果较好，如右图所示，包含了 747 个特征点。

2. 特征点粗匹配

SURF 算法完成对待拼接的两幅图像特征点提取之后，建立其两个特征点的集合 M 和 N，下一步需要对两个集合内的特征点建立一对一的配对关系。由于

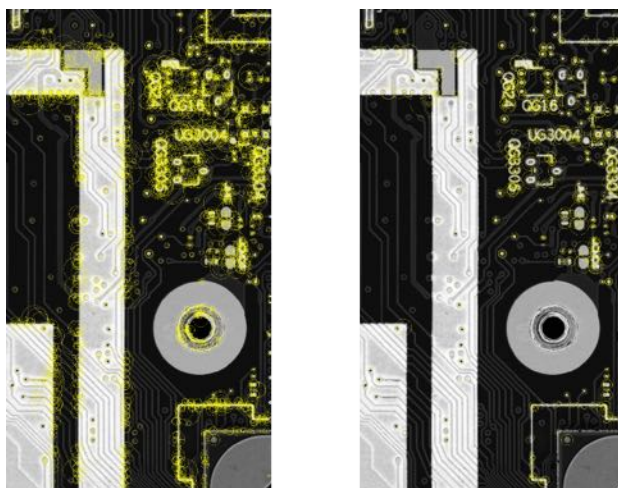


图 3.6 SURF 特征点

刚开始检测到的特征点较多，需要对配对关系进行一个较为简单的粗匹配。一般来说，粗匹配主要有两种算法，一种是暴力匹配，另一种是 K 近邻 (K-Nearest Neighbor, KNN)^[53] 匹配。由于暴力匹配消耗的时间较长，本文采用 KNN 算法进行配对。

由于图像拼接中的特征匹配是一一对匹配，也就是将每两点归为一类。而将 KNN 算法的 K 值设为 1 时，就属于极端情况，非常容易发生过拟合。因此，在实际应用中，采用的是一种最近邻近似算法。具体来说，将两幅图像的特征点集合分别看成训练集和查询集，K 值设为 2，遍历训练集中的特征点，对每个点在查询集中寻找其配对的两个点。配对的原则为两点之间的欧氏距离，因此，训练集中的每个特征点配对的将是查询集中与其欧氏距离最近和次近的特征点，距离分别为 d_0 和 d_1 。然后对距离比 r 进行评估：

$$r = \frac{d_0}{d_1} \quad (3.1)$$

如果 r 的值小于某个阈值，就保留该对匹配点，反之，则抛弃。值得注意的是，阈值的设定是非常重要的，对匹配结果的影响巨大。由于一对好的匹配点应当是最优匹配的，也就是说，对于训练集中的某一点，其应当只与查询集中的最优匹配点距离近，而与其他点距离远。因此， r 的值越小，说明匹配度越高。SIFT 的作者 Lowe 通过大量实验，给定阈值的范围是 0.4-0.6，本文选取一般情况下的 0.5。下面的图 3.7 展示了该算法在两张 PCB 图像上的匹配关系，理论上所有匹配点的连线应当是水平的，而图中仍存在着少量明显的斜线。因此，需要进一步对匹配对进行筛选。

3. 基于主要范围的匹配对细筛选

为了进一步提高特征点匹配的准确率，需要在 KNN 粗匹配之后再次细筛，剔除错误的匹配对。经典的方法是采用随机一致性采样 (Random Sample Consensus,

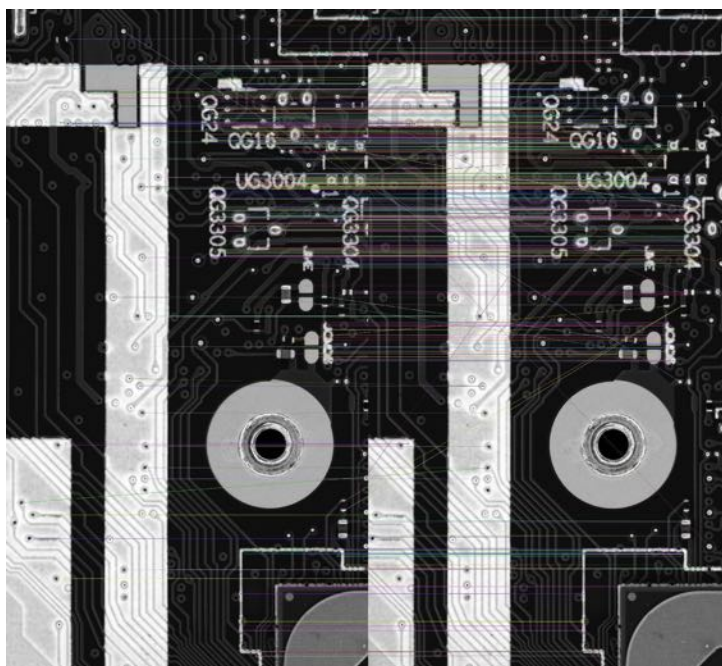


图 3.7 KNN 匹配结果

RANSAC)^[54]进一步对匹配特征点进行提纯。然而，如果粗匹配得到的匹配关系错误过多，也会严重影响 RANSAC 筛选的效率。

在 PCB 图像拼接的特定场景下，有不少同种型号的元器件同时存在于同一张板子的不同区域，而粗匹配很有可能将元器件的相似特征误认为是重叠区域的匹配特征。因此，为了解决该问题，本文提出一种基于主要范围的匹配对筛选方法。由于 PCB 重叠主要是因为平移变换造成的，而偏移距离也可以经测量得出大概的变化范围，因此，可以得出结论，如果将两幅图像直接拼接在一起，那么匹配点之间的线段长度就是偏移距离。我们对所有线段长度进行统计，结果如 3.8 所示：

可以看出，绝大部分的线段长度分布在一个较稳定的范围内，我们将此范围定义为主要范围，而错误匹配的长度偏差较大。上图的统计结果主要范围为 1800~1900，这与实际测量的数据大致吻合。因此，我们将其他不在该范围内的匹配对剔除，图 3.9 展示了经过主要范围筛选法之后的匹配对结果。

可以明显地观察到，图中已经没有了斜线，可见经过主要范围法筛选后的结果已经将匹配错误较大的匹配对点剔除。

4. 图像融合

图像融合用于将两张待拼接图像融合成一张完整的图像。在上述步骤的基础上，用 RANSAC 算法对细筛选后的匹配对进一步提纯，以得到准确的单应矩阵。RANSAC 是一种鲁棒估计方法，能够在存在离群值（例如错误匹配点）的情况下仍然获得相对准确的估计。在 OPENCV 中，它被包装在函数 `findHomography` 中，用于求出单应矩阵。单应矩阵用于对图像进行透视变换，其目的是改变图像

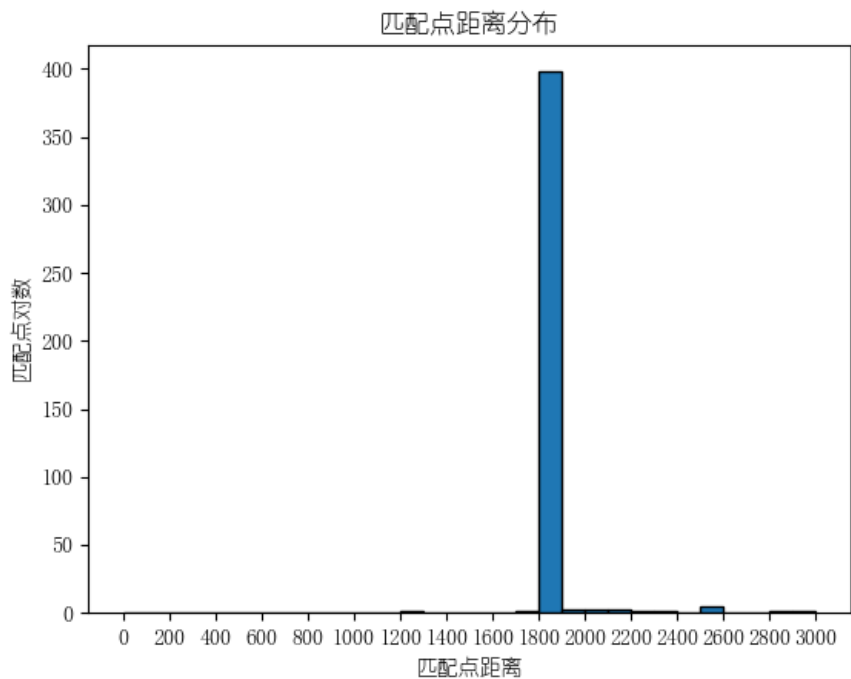


图 3.8 匹配点距离分布

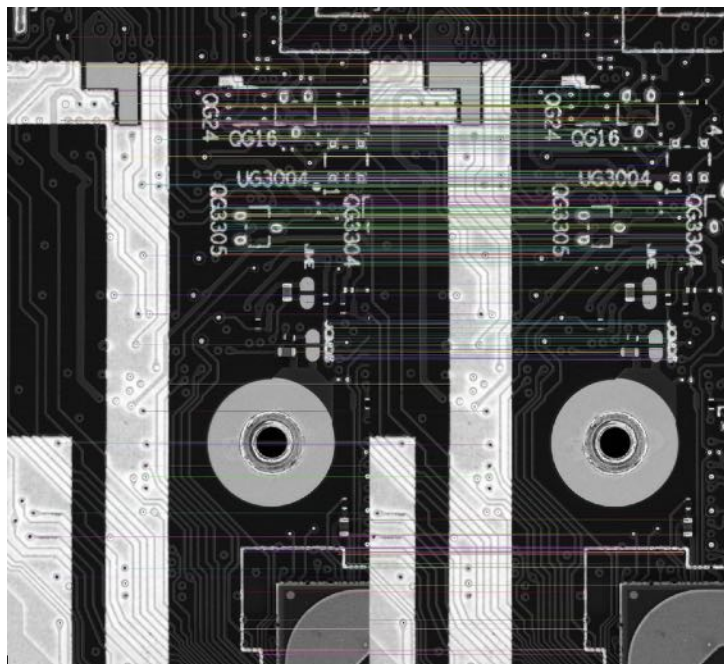


图 3.9 基于主要范围的筛选法处理结果

的视角，视角也就是说图像所处的坐标系。具体来说，在图像拼接的过程中，将两幅图像分别视为参考图和待拼接图像；由于两幅图像间存在平移、旋转等变化导致视角发生变化，通过透视变换将待拼接图像的视角转变为参考图相同的视角；接着构建一个足够容纳两张图像大小的空白底板，先将参考图像映射到对应区域，再将变换后的待拼接图像叠加到底板上，其中重叠区域像素直接覆盖。

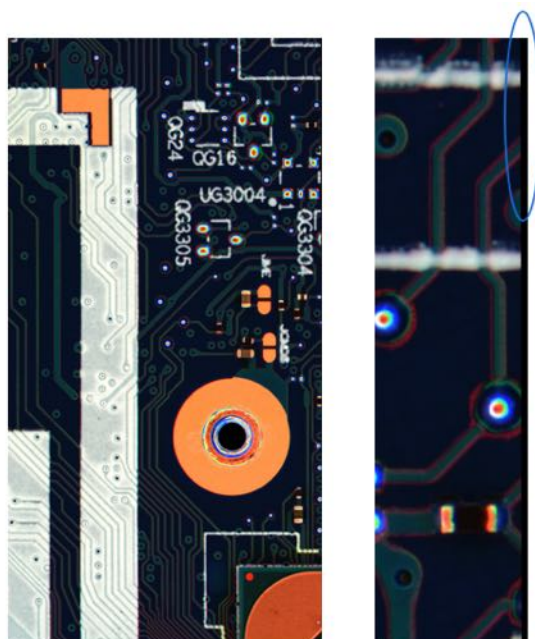


图 3.10 SURF 拼接图像

图 3.10 展示了图像融合之后的拼接图像。左图是拼接图像，右图是拼接图像右上方放大后的视图，不难看出，右上方图像仍存在着少量的黑边。其原因在于待拼接的图像存在微小的旋转变换，通过透视变换的方法无法得到一个完全为矩形的拼接结果，这在下一步拼接时将造成极大地影响。因此，直接采用软拼接算法对 PCB 图像拼接无法适用于本文的场景，下一步将提出新的方法来解决此问题。

3.4.4 基于 SURF 主要距离的主板图像拼接算法

根据前文可知，硬拼接算法具有非常快的速度，但却主要存在以下两点不足：

- (1) 对于不同的拼接对象，偏移量是变化的，无法实时地做出调整。
- (2) 将两幅图像直接拼接后，由于光照不均匀的问题，拼接缝两边会存在较大的色差。

软拼接也存在着难以解决的问题，尽管对特征点的提取匹配等步骤进行了多项优化，但最终在图像融合时仍然会出现拼接结果留有黑边的情况，这使得 PCB 图像整体拼接的流程无法继续下去。

考虑到本文中 PCB 图像具有较强的稳定性，在图像融合时结合硬拼接与软拼接的方法，提出一种基于 SURF 主要距离的主板图像拼接算法。我们弃用软拼接时的图像融合方法，且保留此步骤前的一切改进方案，包括基于划分区域特征点提取和基于主要范围的匹配对细筛选方法，并提出以下两点分别针对硬拼接的两项不足之处进行改进。第一点，在每次拼接相邻图像时计算出 SURF 匹配对的主要距离，并将其作为偏移量以实时调整；第二点，通过基于空域的距离加权融合方法对拼接缝周围的像素重新调整，以消除拼接色差。

1. 基于 SURF 匹配对的主要距离法

对图 3.8 得到的结果进一步分析，将主要范围精确到像素单位，得到的结果如图 3.11 所示。统计结果细化后可以观察到精确到像素单位的匹配点距离分布，理论上相邻图像间的偏移距离应当是唯一的，因此直接将匹配距离中的众数作为偏移量，即主要距离。通过在硬拼接的流程中，每次用 SURF 算法计算两张待拼接图像间的匹配对的主要距离作为偏移量。该方法用算法的手段避免了手工测量带来的误差，并且能够根据不同的拼接对象做到实时调整，很好地弥补了硬拼接存在的不足。

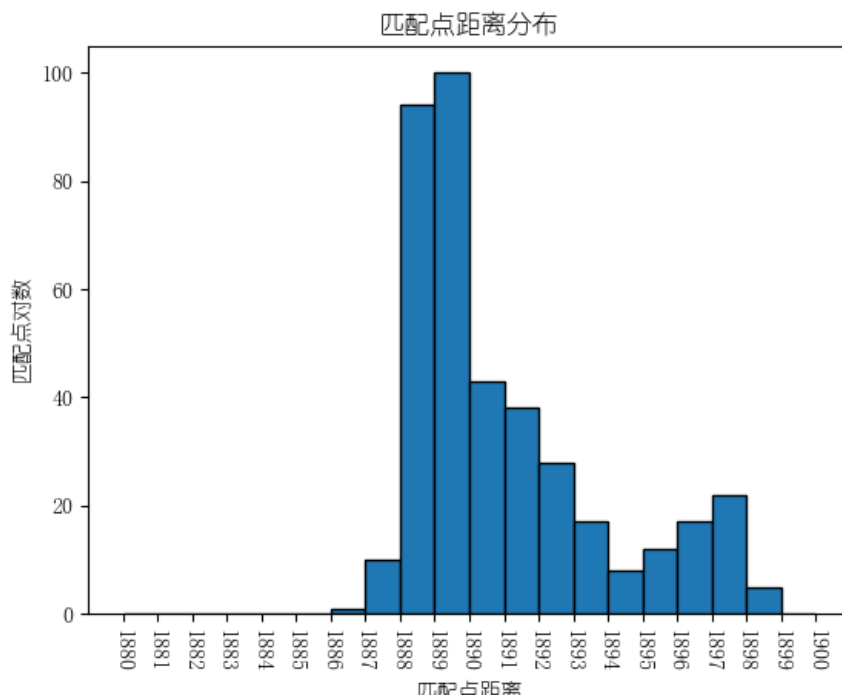


图 3.11 主要匹配点距离分布

2. 基于空域的像素加权融合法

空域图像融合^[55]是指在空域中对原始像素按照某个权重比例进行融合，通常不具备分解转换等过程。光照不均匀导致不同时刻下拍摄的 PCB 图像之间存在着色差，这在图像拼接的拼接缝处非常明显，需要对其处理以平衡两边的色

差,使整张图像更加自然。通过将拼接缝两边一定范围内的像素分别用两张图像的原像素加权得到,在空域上进行图像融合,能够很好地解决该问题。假设待拼接的两张图像分别为 A 和 B, I_1 和 I_2 分别代表 A 和 B 中的像素点,则像素加权法的数学过程为:

$$I(x, y) = \begin{cases} I_1(x, y) & (x, y) \in A, A \cap B = \emptyset \\ \omega I_1(x, y) + (1 - \omega) I_2(x, y) & (x, y) \in A \cap B, A \cap B \neq \emptyset \\ I_2(x, y) & (x, y) \in B, A \cap B = \emptyset \end{cases} \quad (3.2)$$

其中, ω 为加权系数,取值范围为 $0 \leq \omega \leq 1$ 。

3. 列拼接偏移量补偿

由于本文 AOI 系统的在采集 PCB 图像时采用逐列拍摄路线,拼接算法需要先拼接所有同列的局部图像,再将所有列拼接成完整的主板图像。在实际应用中,主要距离法在计算不同相邻图像间的偏移时各有差异,这就导致不同列的图像拼接完成后尺寸不一致,无法进行后续的拼接流程。因此,统一列拼接图像的尺寸是非常重要的。我们采用偏移补偿的方法,将第一列拼接完成后的偏移量之和作为基准,后续的列拼接整体偏移量需保持一致,以保证所有列拼接图像有相同的尺寸。

3.5 实验设计与分析

3.5.1 实验数据

实验所使用的数据全部来自本文 AOI 系统在实际产线中采集得到的 PCB 图像,拍摄规则是按 9 行 7 列划分 PCB,相邻图像间均有一定的重叠,每张图像的尺寸为 3648x5472。图 3.12 展示了部分图像,其中图 3.12(a) 与图 3.12(b) 在垂直方向上有重叠,图 3.12(a) 与图 3.12(c) 在水平方向上有重叠。另外,本章方法的目的在于将 AOI 系统拍摄的多张 PCB 局部图像拼接成整板图像,这需要通过多次拼接不同位置的相邻图像来组成最终的结果。因此,又分别设置了相邻图像拼接实验来验证本文方法在每次拼接操作时的优越性和 PCB 图像整体拼接实验来验证本文方法在工业场景中的实用性。具体内容如下:

(1) 在相邻图像拼接实验中,采用 6 组各不相同的相邻图像作为实验数据,特别地,有 3 组是垂直方向相邻,另外 3 组是水平方向相邻。基于此又设置了对比实验和消融实验,其中对比实验通过将本文方法与 SURF 拼接算法进行成像质量的对比,以验证其有效性;消融实验分别对两项改进方案,包括基于划分区域的特征点提取方案和基于主要范围的匹配对细筛选方案,以验证各自的有效性。

(2) 在 PCB 图像整体拼接实验中, 采用 5 种不同型号主板拍摄的图像作为实验数据, 并以拼接耗费时间为评价标准, 验证方法的实用性。

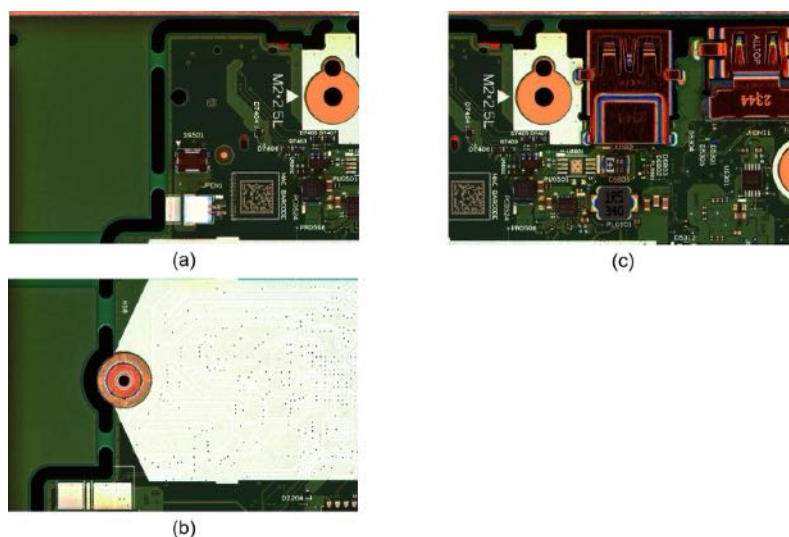


图 3.12 拼接实验数据部分示例

3.5.2 实验环境

本文算法全部在同一个环境中进行实验, 软硬件的关键配置参数如表 3.2 所示。

表 3.2 拼接实验环境配置参数

类型	型号	参数
操作系统	Ubuntu	20.04
GPU	NVIDIA Geforce RTX3090	24GB
CPU	Intel(R) Core(TM) i7-11700K	8 核 16 线程
内存	-	32GB
Python	-	3.6.2
opencv-contrib	-	3.4.2.16

3.5.3 评价指标

针对 AOI 系统下 PCB 图像拼接算法的性能评估, 我们主要从两个方面出发。首先是对于拼接成像的质量评估, 采用均方误差法 (Mean Square Error, MSE) 和结构相似性度量 (Structure Similarity Index Measure, SSIM); 其次是对于拼接速度, 用拼接完成耗费的时间进行评价。

MSE, 表示两幅图像的差异程度。其计算过程是将两幅图像逐像素地作差并平方, 最后取累积和的均值。MSE 的值越小, 说明两幅图像的相似度越高。其公式如下:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (3.3)$$

其中, I 和 K 分别代表两幅图像。

$SSIM$, 用于衡量两幅图像的相似程度。相比于 MSE 只考虑像素间的绝对误差不同, $SSIM$ 是一种感知模型, 更贴近于人眼的直观感受。 $SSIM$ 从三个特征出发对图像进行综合评价, 包括亮度、对比度和结构。用 x 和 y 分别表示两幅待测图像, 其计算过程如下:

对于亮度, 将图像转为灰度图后, 像素值就代表着灰度值, 对于整张图像, 用平均灰度值来表示图像亮度:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (3.4)$$

用 $l(x, y)$ 来表示两幅图像的亮度对比函数:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (3.5)$$

对于对比度, 在计算完亮度的前提下, 用所有点的灰度与亮度的标准差来估计:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2} \quad (3.6)$$

其对比函数用 $c(x, y)$ 来表示:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (3.7)$$

对于结构的比较, 采用亮度与对比度结合的方式, 用归一化后的 $(x - \mu_x)/\sigma_x$ 来表示, 结构对比函数 $s(x, y)$ 为

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (3.8)$$

其中, σ_{xy} 为协方差:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (3.9)$$

对于上述各对比函数中的中的常量 C_1 、 C_2 和 C_3 , 它们的目的是避免相应的函数值在接近 0 时出现的不稳定情况。最终 $SSIM$ 的表达式是三者的结合:

$$SSIM(x, y) = l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma \quad (3.10)$$

其中, α 、 β 和 γ 决定了不同特征在最终结果中的占比。 $SSIM$ 的取值范围为 $[0, 1]$, 其值越大, 表示两幅图像的相似度越高。在实际使用中, 通常采用滑动窗

口的方式来评估整张图像的质量，即对每个窗口计算 SSIM，再计算加权后的平均值：

$$MSSIM(x, y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \quad (3.11)$$

其中，M 为窗口数量， x_j 和 y_j 分别表示两幅图像中第 j 块窗口中的数据。

3.5.4 实验结果分析

1. 相邻图像拼接实验

(1) 对比实验

在本小节中，将对基于 SURF 的图像拼接与本文提出的拼接算法（ours）进行对比实验，并以手工测量的硬拼接生成图像为标准图像。在 6 组不同型号主板上采集到的相邻局部图像的拼接实验结果见表 3.3，其中前三组是水平方向相邻的图像，后三组是垂直方向相邻的图像。在本实验中，先将两种方法得到的拼接图像转化为灰度图，再分别与基准图像计算 MSE 和 SSIM，以评估两种方法的优劣性。可以观察到，在任何一组的结果中，本文方法得到的 MSE 值和 SSIM 值都优于直接采用 SURF 算法拼接得到的图像，这表明生成的图像质量要更好，且相对来说只需要耗费很少的时间就可以完成拼接过程。另外，由于本文使用的 AOI 系统在采集图像时，设置的列偏移较小，垂直方向相邻的图像所能提取到的特征点也就较少。这就导致在最后两组的实验中，SURF 算法无法完成对两幅相邻图像的拼接。而本文提出的方法，只需要检测出少量的特征点即可完成拼接过程。综上所述，本文方法在拼接效果和稳定性上都有着较好的提升。

表 3.3 相邻图像拼接对比实验结果

		时间 (s)	MSE	SSIM
第 1 组	ours	2.7	81.1	0.16
	surf	30.2	81.2	0.13
第 2 组	ours	2.7	79.3	0.18
	surf	29.2	98.2	0.14
第 3 组	ours	2.7	60.3	0.36
	surf	31.7	61.4	0.32
第 4 组	ours	2.2	17.7	0.91
	surf	67.6	104.1	0.10
第 5 组	ours	2.2	9.9	0.96
	surf	-	-	-
第 6 组	ours	2.2	7.1	0.98
	surf	-	-	-

(2) 消融实验

为了验证本文提出的改进方案的有效性，在本节中，进行了一系列的消融实验，实验数据与对比实验完全相同。以本文实现的最终方法（ours）为基准，将提出的改进方案依次剔除进行实验，包括基于划分区域的特征点提取和基于主

要范围的匹配对细筛选，分别用 DR 和 MS 表示。对于结果的比较仍然采用对比实验中的相同方案。表 3.4 展示了实验结果，可以观察到，本文方法所得的结果都是最佳的。

基于划分区域的特征点提取：-DR 表示只去除基于划分区域的特征点提取方法。从表 3.4 中可以看出，结果全面落后于本文方法。特别是在时间指标上，由于特征点提取区域扩大为整张图像，导致提取的时间明显变长，平均增加了 2s 左右。

基于主要范围的匹配对细筛选：-MS 表示只去除基于主要范围的匹配对细筛选方法。观察表 3.4 可知，第 2、3、4 组的结果与本文方法完全相同，这是因为最终得到的偏移量是相同的。然而，在其他组的结果中，都要比本文方法差，这是由于 PCB 图像中部分贴片元件形状极其相似，比如两个电阻元件，尽管两者不同属于重叠部分，但是 SURF 算法误认为两者属于相同特征，导致在匹配时出现多处错误匹配。

表 3.4 相邻图像拼接消融实验结果

		时间 (s)	MSE	SSIM
第 1 组	ours	2.7	81.1	0.16
	-DR	4.4	117.2	0.12
	-MS	2.7	115.8	0.12
第 2 组	ours	2.7	79.3	0.18
	-DR	4.8	90.8	0.14
	-MS	2.8	79.3	0.18
第 3 组	ours	2.7	60.3	0.36
	-DR	4.3	94.3	0.23
	-MS	2.7	60.3	0.36
第 4 组	ours	2.2	17.7	0.91
	-DR	4.6	39.0	0.90
	-MS	2.2	17.7	0.91
第 5 组	ours	2.2	9.9	0.96
	-DR	4.5	29.0	0.94
	-MS	2.2	53.8	0.78
第 6 组	ours	2.2	7.1	0.98
	-DR	4.9	33.1	0.87
	-MS	2.2	96.6	0.49

2. PCB 整板图像拼接实验

拼接算法的最终目的是实现单张主板图像的完整拼接，并且需要保证足够的拼接速度与成像质量。在本小节中，将对 AOI 系统采集的 5 种不同型号的主板图像进行拼接实验。每张完整的主板图像分为 9 行 7 列共 63 张局部小图，每张拍摄的小图都是尺寸为 3648×5472 的 RGB 图像。拼接时采用先拼接所有列，再将所有列拼接成完整图像的策略，表 3.4 记录了 5 种主板图像拼接完成所耗费的时间。

可以得到，单张 PCB 整板拼接完成的时间约为 32s，这包括了图像读取时

表 3.5 PCB 整板图像整体拼接实验结果

型号	读取时间 (s)	拼接时间 (s)	存储时间 (s)	总时间 (s)
主板 1	15.57	6.12	10.75	32.44
主板 2	15.53	6.10	10.59	32.22
主板 3	15.51	6.13	10.62	32.26
主板 4	15.48	6.13	10.56	32.17
主板 5	15.49	6.15	10.64	32.28

间、图像拼接时间和图像写入硬盘的时间。由于拼接过程中采用的是高分辨率的原图像，无论是单张图像还是最后的结果图尺寸都比较大，因此图像的读取和写入耗费了大量的时间。不过，由于图像拼接是在缺陷检测的过程中就启动的，在检测时已经对图像做了读取操作，因此在实际应用中，该部分时间可以省略。最终得到图像拼接实际耗费的时间在 16s 和 17s 之间。

图 3.13 是主板 1 拼接完成后的结果图，其尺寸为 32362×26916，为后续的小目标缺陷可视化工作提供了良好条件。

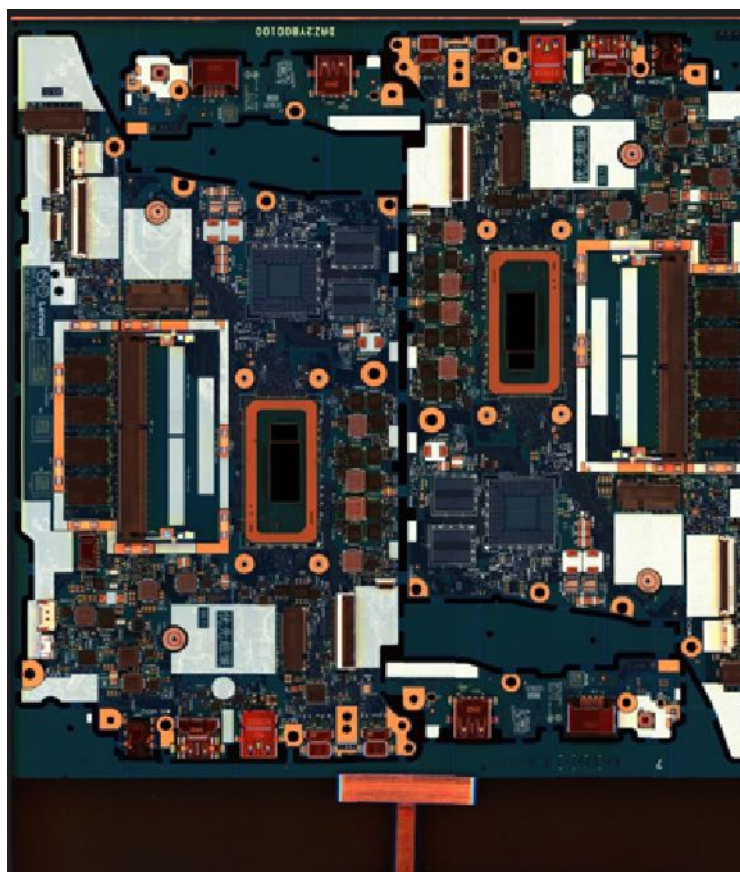


图 3.13 主板 1 拼接成像图

3.6 本章小结

本章针对 AOI 系统采集的多张局部小图像拼接成完整大图的问题进行了研究，基于本文所依赖的 AOI 系统，对图像拼接的基本规则和流程都进行了优

化,充分结合硬拼接和软拼接的优势,提出一种基于 SURF 主要距离的主板图像拼接算法。具体来说,采用硬拼接作为相邻图像间的拼接方式,有效地保证了图像拼接的速度和稳定性;提出以 SURF 特征点匹配的方式建立两张待拼接图像间的联系,并且提出基于划分区域的特征点提取方法和基于主要范围的匹配对细筛选方法对 SURF 算法进行改进,以 SURF 特征点匹配对的主要距离作为相邻图像间的偏移距离,这使得硬拼接流程中无法处理不同位置处相邻图像间的偏移量不同的问题得到了很好地解决;使用基于空域的像素加权融合法消除拼接后图像在拼接缝处产生的色差,以尽可能地消除拼接裂缝;使用偏移补偿的方法,保证所有列拼接图像尺寸相同。实验表明,本文方法不仅在两幅图像拼接中优于 SURF 软拼接,在整板拼接时,也依然能够保证高分辨率大图的高效拼接,较好地解决了以上述的一系列问题。

第4章 基于自适应协同特征的 DETR 缺陷检测算法研究

4.1 引言

在 SMT 的生产流程中,贴片后的 PCB 需要经过回流焊环节使得焊盘和元件充分结合,焊接时的高温在融化锡膏的同时会使得部分贴片元件位置发生不同程度的变化,这就容易出现缺陷情况,如元件的偏移、立碑、侧立等。这就需要对缺陷情况进行检测,以避免次品进入后续流程。缺陷检测算法用于解决此问题,它是整个 AOI 系统中最核心的算法,是质量检测最核心的任务,不管是 AOI 系统平台的设计,还是上一章中拼接算法的研究,都直接或间接地服务于缺陷检测。具体来说,在大部分的 AOI 系统中,软件平台是算法的载体,而第 3 章介绍的拼接算法不仅用于检测前的区域标定,也用于检测后的整板可视化工作。

贴片元件的缺陷检测的主要有两个核心问题:缺陷定位和缺陷识别。当前常用的方法是通过模板匹配算法,将拼接后的 PCB 图像与标准模板比对,先确定待检测区域的位置,然后利用设定好的判断准则来对当前元件进行缺陷判定和分类。尽管该算法相对于人工目检有着很大的效率提升,但其本身仍然存在着较大的不足。一方面,该算法对模板质量要求非常高,这需要专业人员耗费大量时间来进行繁琐的设计流程,且匹配时的判定规则也比较单一,难以应对复杂的缺陷特征,使得误检率较高。另一方面,模板匹配严重依赖于前置的图像拼接任务,拼接成像的质量也在很大程度上决定着检测的精度。

基于深度学习的目标检测技术主要目的就在于目标物体的定位和识别,这与贴片元件缺陷检测的核心问题相契合。将目标检测技术应用于 SMT 贴片元件的缺陷检测任务中,优势在于:算法不再受制于前处理阶段,将拍摄的图像直接作为模型的输入数据,图像拼接只用于最终的缺陷可视化工作,实现了端到端的检测流程。然而,不同于一些常见的工业缺陷检测问题,SMT 场景中仍然存在着多种复杂问题,以至于目前还没有一种目标检测方法实际应用于该领域。究其原因,首先,在实际生产中,可标注的 PCB 图像数据数量稀少,也没有任何相关的公开数据集供学术界研究;其次,贴片元件本身特征复杂,绝大部分贴片元件本身非常微小且存在聚集现象;而且,由于相机视场的限制,一张 PCB 图像是分成若干张局部图像分别检测的,这对算法的实时处理能力要求很高。

针对解决以上问题,本章提出一种基于自适应协同特征的 DETR 缺陷检测算法。具体来说,主要对以下内容进行了研究:(1)针对现实中负样本稀少的情况,采用多种方法收集与制作数据,建立了本文所使用的数据集;(2)针对工业场景中对检测精度和实时性的要求,采用轻量化的实时目标检测框架 RT-DETR 作为的基准模型,并在此基础上进一步地优化;(3)针对数据集中不同缺陷的形

状特征，提出自适应协同模块 ACBlock (Adaptive collaboration Block)，提升模型对于形状不规则目标的特征学习能力；(4) 由于缺陷元件在图像中占比很小，属于小目标的范畴，设计 IMIoU (Inner and MPD IoU) 损失函数，可以更有效地捕捉到微小目标信息，并且具有较快的收敛速度。

本章结构安排如下，4.2 节对 RT-DETR 目标检测框架作出了具体说明。4.3 节设计了一种用于 SMT 贴片元件缺陷检测的算法，并对该方法进行了详细介绍，包括基准模型的结构、特征提取模块和损失函数。4.4 节是实验结果与分析，首先介绍了本文提出的 SMT 贴片元件缺陷数据集和一些公开数据集，然后将本文提出的检测算法在多个数据集上开展实验，包括与其他模型的对比实验和本身的消融实验。4.5 节给出了本章的总结。

4.2 RT-DETR

DETR 系列算法经过一系列的发展，已然成为目标检测领域最杰出的算法，在各大公开数据集的榜单上都达到了 SOTA。这不仅在于 DETR 本身优秀的架构，也受益与当前大模型时代涌现出的各种预训练策略。在提升模型精度的同时，也使得模型尺度变得臃肿，不利于工业应用。

百度提出的 RT-DETR 很好地解决了这个问题，它在 DETR 的整体架构上做出了极大地优化，真正为工业界提供一款了实用性较高的 DETR 系列的实时检测器。在 RT-DETR 之前，DETR 系列模型采用纯粹的 transformer 处理多尺度特征图。需要说明的是，transformer 起源于自然语言处理领域，用于处理序列化的文本序列。因此，在处理更加多维的图像特征时，需要将特征图展平成一个序列再输入。DETR 将多个层级特征图拉平并拼接成一个非常长的序列作为输入，如 Deformable DETR 中使用了 ResNet 的 C_3 、 C_4 和 C_5 层特征图，以处理多尺度信息。为了使多尺度特征图得到充分的交互，不得不采用多层的编码器来处理（通常为 6 层），而编码器的计算复杂度是像素数量的平方级，这造成了极大的计算量。深层特征通常具有更丰富的语义信息，这对区分不同物体非常有用，而浅层特征的语义信息较为匮乏，直接将两者等同处理会造成资源的浪费。

因此，为了将深层特征和浅层特征分开处理，RT-DETR 将多层编码器的替换成两个独立模块，包括一个 AIFI 模块 (Attention-based Intra-scale Feature Interaction) 和一个 CCFM (cross-scale feature-fusion module) 模块。其中 AIFI 即是一层的 transformer 编码器，与 DETR 相比，数量从 6 个减少到了 1 个，且只处理最后一层 (S_5) 的深层特征图，极大地减少了计算量。CCFM 是一个特殊设计的跨尺度特征融合模块，与 FPN 的思想类似。较浅的两层特征图 (S_3 和 S_4) 无需经过处理，直接输入到 CCFM 中，与 AIFI 处理深层特征得到的新特征图 F_5

进行特征融合。这样巧妙的设计，使得由注意力机制主导的多尺度特征融合转变到由卷积层主导，在降低计算量的同时又不失特征的充分交互，可谓两全其美。整体的计算过程如下：

$$\begin{aligned}
 \mathbf{Q} &= \mathbf{K} = \mathbf{V} = \text{Flatten}(S_5), \\
 F_5 &= \text{Reshape}(\text{Attn}(\mathbf{Q}, \mathbf{K}, \mathbf{V})), \\
 \text{Output} &= \text{CCFM}(\{S_3, S_4, F_5\})
 \end{aligned} \tag{4.1}$$

其中， \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 是都注意力机制中的计算量，分别表示查询向量、键向量和值向量，它们都是由输入的特征向量经过三个不同的全连接层得到。在 RT-DETR 中，注意力机制只用于处理 S_5 的特征图，因此，需要将 S_5 的特征展平后输入，再将处理后的特征再重塑为三维的图像特征图 F_5 ，以便于同其他两个层级的特征图做进一步处理。

另外一点，DETR 中查询选择的作用是从编码器输出的特征序列中选择固定数量的特征作为目标查询，其经过解码器后由预测头映射为置信度和边界框。由 DINO 提出的混合查询选择策略可以看出，优质的查询选择对模型也更加有利。现有的 DETR 系列模型都是利用这些特征的分类分数直接选择得分最高的特征。然而，由于分类分数和 IoU 分数的分布存在不一致，分类得分高的预测框并不一定是和真实框（Ground Truth, GT）最接近的框，这导致高分类分数低 IoU 的框会被选中，而低分类分数高 IoU 的框会被丢弃，这将会损害检测器的性能。RT-DETR 的解决方案是提出了 Iou-aware 查询选择方法。该方法在训练期间对检测器施加约束，使得高 IoU 的特征产生高分类分数，低 IoU 的特征产生低分类分数。这样一来，模型在选择分类分数较高的特征时，对应的预测框能够同时具有高分类分数和高 IoU 分数，下式阐述了该方法的原理。

$$\begin{aligned}
 \mathcal{L}(\hat{y}, y) &= \mathcal{L}_{\text{box}}(\hat{b}, b) + \mathcal{L}_{\text{cls}}(\hat{c}, \hat{b}, y, b) \\
 &= \mathcal{L}_{\text{box}}(\hat{b}, b) + \mathcal{L}_{\text{cls}}(\hat{c}, c, \text{IoU})
 \end{aligned} \tag{4.2}$$

其中， \hat{y} 为预测量， y 为真实量， c 表示类别， b 表示边界框， L_{box} 和 L_{cls} 分别表示边界框回归损失和分类损失。

借助各个模块的高效优化，RT-DETR 表现出卓越的性能。在与 FPS 相当的实时目标检测器相比时，实现了检测精度上的超越；在与两阶段算法和大尺度 DETR 系列目标检测器相比时，在速度上也做到了提升。

4.3 改进的 RT-DETR 模型设计

本文将 SMT 贴片元件缺陷检测看成是一种目标检测任务，基于 RT-DETR 设计了一种全新的轻量级 SMT 贴片元件缺陷检测实时算法，并命名为 SMT-DETR，整体结构如图 4.1 所示。在高效实时目标检测器的基础上，针对 SMT 场景中的特别需求，提出了一种全新的自适应协同模块 ACBlock，并结合 MPDIoU^[56] 和 InnerIoU^[57] 设计了全新的 IMIoU，改进后的模型在自制的 SMT 数据集上性能显著提高。同时，本文方法也可以扩展到其他类似的领域，在一些公开数据集上也取得了不错的效果。

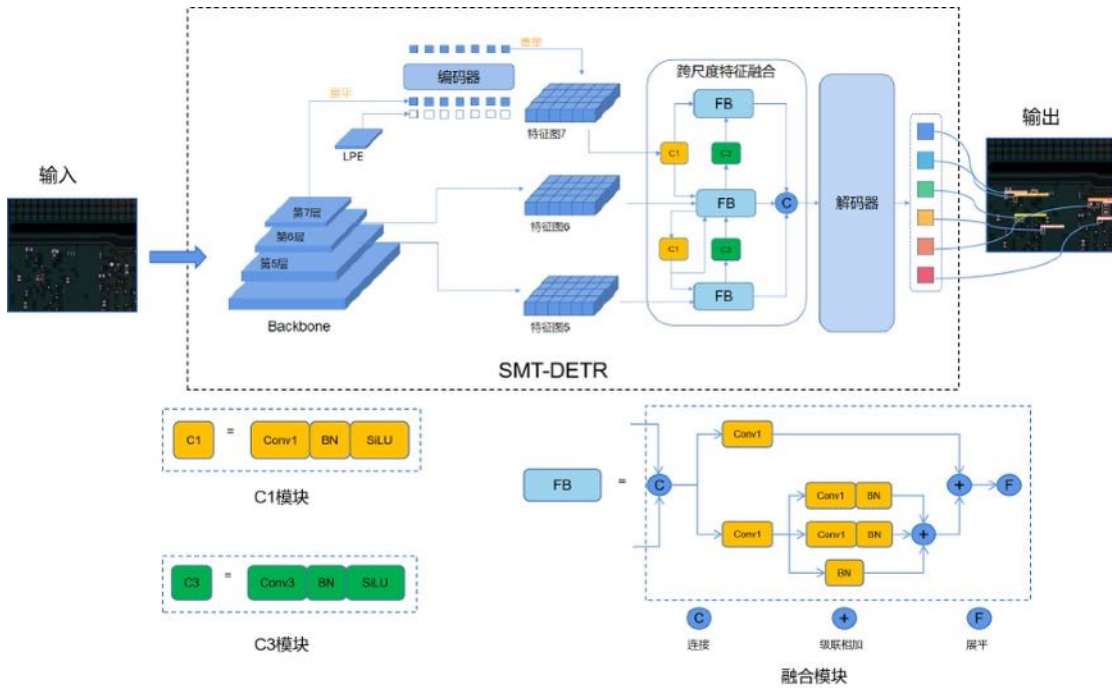


图 4.1 SMT-DETR 结构图

4.3.1 自适应协同模块

在整体架构更先进的同时，也离不开一些细节模块的性能提升。通道注意力机制旨在寻找一种有损压缩方法以实现最佳的特征表示，它通常被插入到网络中作为一个增强特征的模块。OrthoNets^[58] 引入了 SENet^[59] 的一个变体 OrthoAttention。如图 4.2 所示，OrthoAttention 首先将一组尺度为 $(1 \times h \times w)$ 的随机滤波核进行正交得到正交滤波核，再经过挤压和激励操作，得到一个 $(c \times 1 \times 1)$ 的权重向量，最后通过乘法逐通道加权到输入的特征图上。通过这些操作，能够从特征空间的正交子空间中提取信息，使得特征图具有更加丰富的表示。以 $K \in \mathbb{R}^C \times \mathbb{R}^H \times \mathbb{R}^W$ 表示这些滤波器，OrthoAttention 的处理过程可以解释为

$$\mathbf{F}_{\text{ortho}}(X)_c = \sum_{h=1}^H \sum_{w=1}^W K_{c,h,w} X_{c,h,w} \quad (4.3)$$

其中, c 代表通道数, h, w 分别代表长和宽。

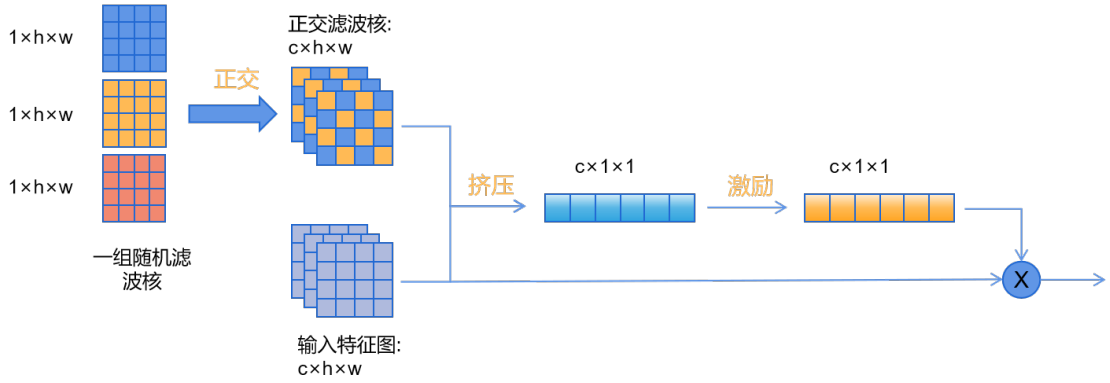


图 4.2 OrthoAttention 示意图

另外, DCN^[60-61]通过对卷积核施加偏移量, 改变了传统卷积模块感受野的固定形状, 增强了模型对目标几何变换的建模能力。以 3×3 大小的卷积核, 步长为 1 时普通卷积为例, 对于特征图 $X \in \mathbb{R}^C \times \mathbb{R}^H \times \mathbb{R}^W$ 上的一个位置 p_0 , 需要在以 p_0 为中心周围一圈相邻的 9 个位置 (集合记作 \mathcal{R}) 采样, 作用域是一个标准的矩形 (图 4.3 (a)), 则在 p_0 处的输出 $y(p_0)$ 为:

$$y(p_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} w(\mathbf{p}_n) \cdot x(p_0 + \mathbf{p}_n) \quad (4.4)$$

DCNv1 在每次采样的基础上新增了一个可学习的距离偏移量 Δp_n , 作用域不再是矩形, 而是不规则的多边形 (图 4.3 (b)), 此时 $y(p_0)$ 变为:

$$y(p_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} w(\mathbf{p}_n) \cdot x(p_0 + \mathbf{p}_n + \Delta p_n) \quad (4.5)$$

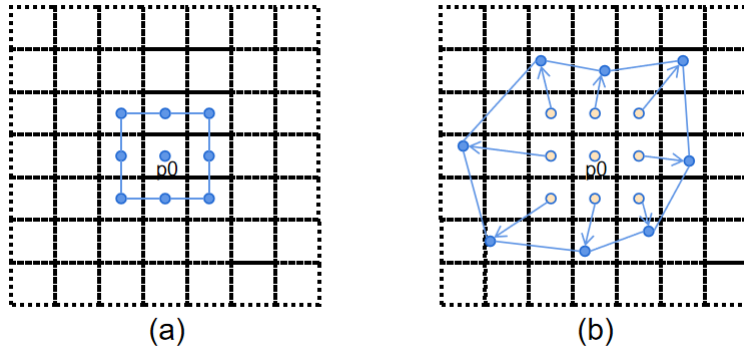


图 4.3 DCN 采样示意图

添加偏移量后, 能够更好地适应具有几何变换性质的目标。然而, 不规则的作用域很可能会覆盖到超出目标区域的不相关内容, 导致特征受到过多干扰。DCNv2 在每次计算采样点信息时加入权重项 Δm_n , 通过该可学习的参数来抑制不相关区域的影响, 此时 $y(p_0)$ 变为:

$$y(p_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} w(\mathbf{p}_n) \cdot x(p_0 + \mathbf{p}_n + \Delta p_n) \cdot \Delta m_n \quad (4.6)$$

本文结合两者,构建了一种全新的 ACBlock,由四个部分组成,包括一个 CBS 模块(Conv, Batch Normalization, SiLU),一个 DB 模块(DCN, Batch Normalization),一个通道注意力模块(OrthoAtten)和一个残差模块(Conv1)。ACBlock 的设计思想在于当 DCN 使得特征变得更加复杂难学习时,在其后添加 OrthoAttention 提升网络的学习能力,充分结合了两者的优势,并通过实验验证确定了 ACBlock 模块的最佳放置位置。本文对 Backbone 最后三个层级的网络 {S5, S6, S7} 进行了重构,其中 S5 和 S6 添加 OrthoAttention,分别输出两种不同尺度的特征图 Feature5 和 Feature6。我们将 ACBlock 插入 S7,特征图经过 ACBlock 和 AIFI 得到 Feature7。重构后的 S5 和 S6 结构如图 4.4 所示,S7 如图 4.5 所示。

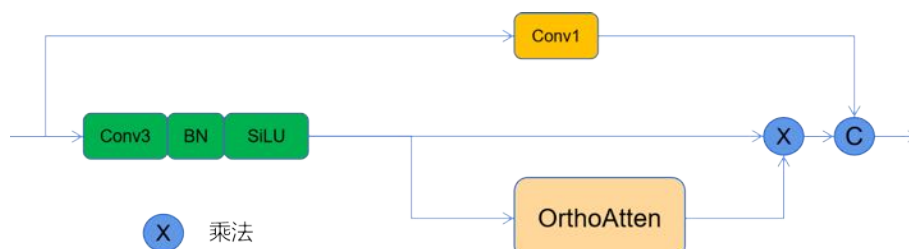


图 4.4 正交注意力模块结构图 (S5、S6)

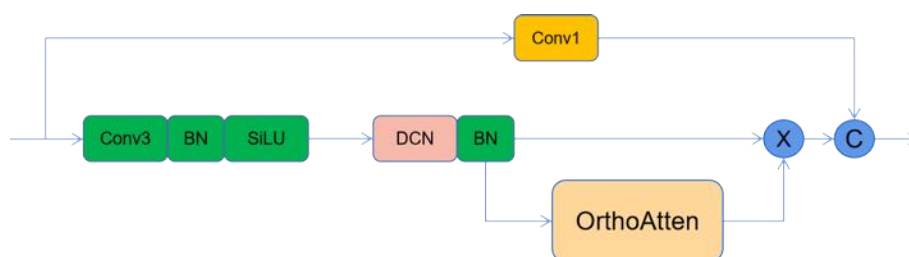


图 4.5 自适应协同模块结构图 (S7)

最后,将原本的正余弦绝对位置编码改造为可学习的位置编码,提升了模型的泛化能力。

4.3.2 损失函数

目标检测的目的是找出图像中所有感兴趣的目标,并获得目标的类别和位置信息,因此,它同时包含了两项基本任务,即分类和位置回归任务,分别对应两种损失:分类损失 L_{cls} 和边界框回归损失 L_{box} 。

1. 分类损失

分类损失用于计算预测目标的类别与实际类别之间的距离,帮助模型学习正确地分类检测到的目标。交叉熵这一概念来自于信息论,可以用于衡量两个分布之间的距离。在深度学习中,将离散的真实标签经过 one-hot 编码后会为每个目标都生成一个二进制向量,而神经网络在经过 softmax 层输出预测物体的类别时也会生成一个概率向量,交叉熵损失用来衡量这两者之间的距离。对于二分类

的交叉熵损失函数，对其的计算公式有如下定义：

$$CE(p, y) = \begin{cases} -\log(p) & \text{if } y = 1, \\ -\log(1 - p) & \text{otherwise.} \end{cases} \quad (4.7)$$

其中， y 为标签，在二分类任务中取值为 0 或 1，分别代表图像中的背景和前景， p 表示预测正确的概率。接下来，用 p_t 代替 p ：

$$p_t = \begin{cases} p & \text{if } y = 1, \\ 1 - p & \text{otherwise.} \end{cases} \quad (4.8)$$

这样交叉熵损失就可以简化为：

$$CE(p, y) = CE(p_t) = -\log(p_t) \quad (4.9)$$

由于工业场景中正负样本非常不平衡，在大多数的检测图片中，只存在极少的缺陷情况，即正样本数量远小于负样本。如果继续用交叉熵损失，就可能导致过拟合，模型更容易判断为负样本，而对正样本的识别能力变差。FocalLoss 能够较好地解决这个问题，它可以被表示为：

$$FL(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t) \quad (4.10)$$

其中， α_t 为权重因子，用来平衡正负样本损失，范围在 $[0, 1]$ ，其大小根据正负样本的比例设置。 γ 为调节因子，范围在 $[0, 5]$ ，当 γ 为 0 时，就变为了最普通的交叉熵损失函数。当 p_t 趋向于 1 时，说明为易分样本，反之为难分样本，两者得到的损失值分别被缩小和放大。引入 γ 可以减低易分样本的损失贡献，从而增加难分样本的损失比例。从直观上来说，比例较大的样本即为易分样本，较小的即为难分样本。因此，FocalLoss 不仅使得模型的分分类准确度提高，还很好地缓解了样本不平衡的问题。

2. 边界框回归损失

目标检测任务中的边界框回归损失一般采用基于 IoU (Interaction over Union) 的损失，IoU 即是模型预测的边界框 (bounding box) 与真值框 (Ground Truth box) 之间的交并比，取值范围为 $[0, 1]$ ，值越大说明预测得越准确，它可以被表示为：

$$IoU = \frac{\text{Inter}}{\text{Union}} \quad (4.11)$$

然而，IoU 的计算太过简单，它忽略了两个重要的特殊情况，一是两个框无重叠时，IoU 的值始终为 0，也就无法优化；二是 IoU 只关注面积比，导致相同的 IoU 值也可以对应着不同的重叠情况。因此，对 IoU 的改进是非常必要的。

GIoU^[62]引入 bounding box 和 GT box 的最小外接矩形，其面积用 A_c 表示。GIoU 的计算公式如下，它的值在两者完全重合时达到最大，在无重叠时随着框的距离而无线趋向于-1，因此可以很好地反应距离度量。

$$GIoU = IoU - \frac{|A_c - U|}{|A_c|} \quad (4.12)$$

DIoU^[63]从两个框中心点之间的距离出发，用 b 和 b^{gt} 分别代表 bounding box 和 GT box 的中心点， ρ 表示两个中心点间的欧式距离， c 表示最小外接矩形的对角线长度，计算公式如式 4.13。当两个框距离较远时，DIoU 通过距离比值的方法，很好地解决了 GIoU 此时损失值非常大的情况，加快了收敛速度。

$$DIoU = IoU - \frac{\rho^2(b, b^{gt})}{c^2} \quad (4.13)$$

CIoU^[63]在 DIoU 的基础上，增加了一个特殊情况的考虑：当两个框的中心点重合，但宽高比不同时，DIoU 退化为普通的 IoU。其计算公式如下：

$$CIoU = IoU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \quad (4.14)$$

其中， v 是两个框的宽高比的一致性度量， α 是权重因子。

InnerIoU 提出使用辅助边界框来加速收敛过程，并且通过引入尺度因子比 α 来控制辅助边界框的大小。分别给原本的 bounding box 和 GT box 的设置一个辅助框，经过试验发现，对于 IoU 较高的样本，较小的辅助框能够提升 IoU 损失函数梯度的绝对值，对于 IoU 较低的样本，较大的辅助框能够提升 IoU 损失函数梯度的绝对值。因此，通过调节 α 的大小，可以同时提升高 IoU 样本和低 IoU 样本的回归速度，从而加速模型收敛。其计算过程仍与 IoU 相似，如式 4.15 所示。特别地，引入辅助边界框后，对 Inter 和 Union 的计算发生了变化，对于小目标而言，辅助边界框通常为真实框的放大情况，这使得 InnerIoU 在处理小目标问题时非常有效。

$$IoU^{inner} = \frac{Inter}{Union} \quad (4.15)$$

MPDIoU 针对 bounding box 和 GT box 宽高比相同但值不同的情况进一步研究，采用一种基于最小点距离的边界框相似度比较度量，直接在训练过程中最小化两个框的左上角距离和右下角距离。其计算过程如下：

$$MPDIoU = IoU - \frac{d_1^2}{h^2 + w^2} - \frac{d_2^2}{h^2 + w^2} \quad (4.16)$$

其中 d_1 和 d_2 分别代表两个框的左上角距离和右下角距离, $(x_1^{prd}, y_1^{prd}, x_2^{prd}, y_2^{prd})$ 为预测框, $(x_1^{gt}, y_1^{gt}, x_2^{gt}, y_2^{gt})$ 为 GT 框, h 为图片的高度, w 为图片的宽度。通过计算过程可以发现, MPDIoU 在计算 IoU 之外, 只需要计算 d_1 和 d_2 的值, 相比于前面提到的 DIoU 的处理方式, 要更为简便, 也使得 MPDIoU 在全面考虑到影响 IoU 的因素时, 能够获得最快的收敛速度。

考虑到 PCBA 图像中贴片元件普遍较为微小, 提升模型的小目标检测能力是非常必要的。综合以上对一些 IoU 损失函数的论述, 可见 InnerIoU 和 MPDIoU 的优势与本文任务非常契合, 并且两者能够相得益彰。因此, 本文将 InnerIoU 和 MPDIoU 相结合, 用 IoU^{inner} 代替 MPDIoU 中计算 IoU 的过程, 由此设计出全新的 IMIoU 损失, 计算过程如下:

$$\begin{aligned} \mathcal{L}_{IMIoU} &= 1 - IMIoU \\ &= 1 - IoU^{inner} + \frac{d_1^2}{h^2 + w^2} + \frac{d_2^2}{h^2 + w^2} \end{aligned} \quad (4.17)$$

4.4 实验设计与分析

4.4.1 实验数据

基于深度学习的目标检测算法的效果, 在很大程度上依赖于数据集的好坏。因此, 建立一个良好的数据集是非常必要的, 我们基于本文的 AOI 系统的数据采集功能, 将某知名消费电子制造商企业 SMT 产线上真实的 PCBA 主板作为拍摄对象, 同时考虑到现实中负样本稀少的情况, 利用多种数据增强技术扩充了数据, 最终建立起 SMT 贴片元件缺陷数据集 (SMT Chip Component Defects, SMTCD)。此外, 现有与 PCB 相关的公开数据集, 都只关注于 PCB 裸板的缺陷, 如 PKU-Market-PCB^[64]、FICS PCB^[65] 和 DeepPCB^[66] 等, 与本文研究的贴片元件缺陷不同, 但考虑到同属于工业上 PCB 缺陷检测的范畴, 因此选用其中的一项作为实验数据集, 此处选用 DeepPCB。另外, 为了进一步验证本文方法的优越性, 还在具有相似特征的公开数据集 VisDrone2019^[67] 数据集上进行了实验。下面将对所用的三种数据集做详细介绍。

1. SMTCD

SMTCD 数据集来自于某消费电子制造商, 在 SMT 产线上实际生产过程中贴片元件经过回流焊之后, 通过 AOI 系统拍摄的 PCBA 局部图像, 并且还通过数据增强的方法扩充了数据集。数据集的部分示例如图 4.6 所示。

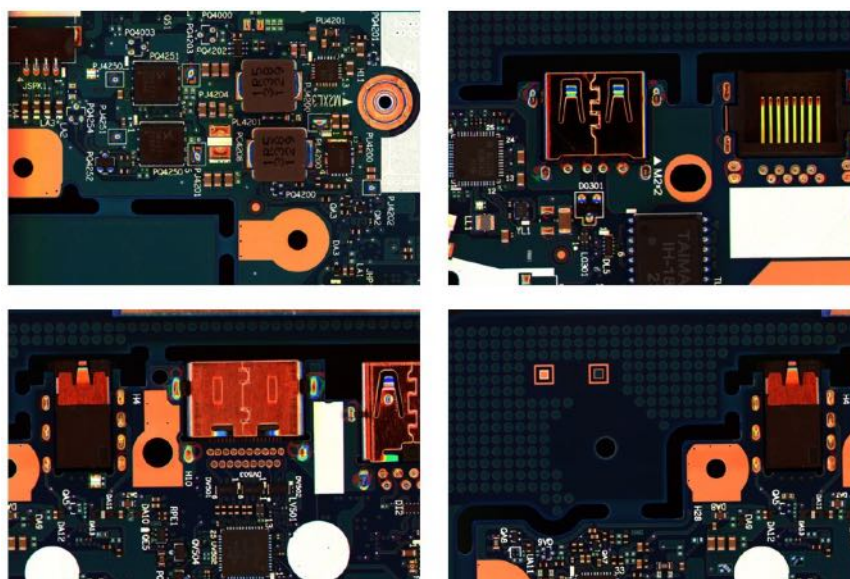


图 4.6 SMTC 数据集部分示例

(1) 原始数据采集与分析

在实际采集数据的过程中，我们能获得两种数据。一种是 PCBA 经过本文 AOI 系统时拍摄的多张局部图像，经过标注后可以直接作为本文算法的输入数据。如图 4.7 所示，图中是 AOI 系统实际采集的数据，第一行和第二行分别是不同型号主板的局部图像，贴片元件包括电阻、电容等作为待检测的对象，按照设计好的 PCB 走线规则分布在主板的任意位置。在很多情况下，贴片元件的分布较为集中，存在着聚集现象，并且绝大多数都非常微小，在原图中只占据很小的一部分。原图的尺寸为 3648×5472 ，单个贴片元件的长和宽尺寸基本在 $[50, 150]$ 之间。

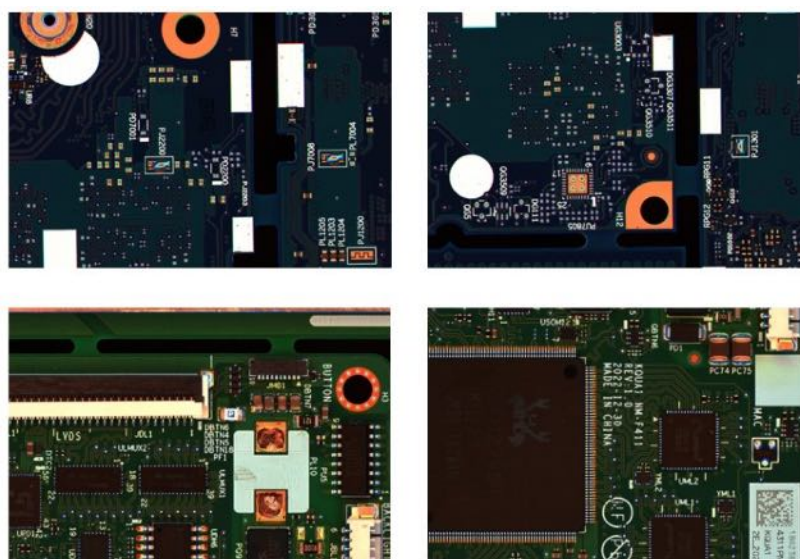


图 4.7 AOI 系统实际采集数据

另一种是原工厂中 AOI 系统检测出来的单一元件缺陷图，类似于目标检测

算法预测的 bounding box 图像，但由于原 AOI 系统的处理方式不同，在图中仍存在着较大冗余面积的背景。由于该工厂生产规模较大，经过一段时间的收集，得到了大量的真实单个元件缺陷图，经过归纳整合后，将收集到的元件图分为 5 种缺陷类型，包括偏移、缺件、立碑、侧立和少锡。典型的样本如图 4.8 所示，第一行是正常的情况（图 4.8(a)），第二行是五种缺陷情况。可以看到，缺陷情况的图片是在正常情况下发生不同程度的形变。具体来说，偏移是指元件电极的一端脱离了焊盘（图 4.8(b)），缺件是指焊盘上元件丢失（图 4.8(c)），立碑是指元件在焊盘的一端竖直站立（图 4.8(d)），侧立是指元件垂直 90 度翻转（图 4.8(e)），少锡是指锡膏过少导致焊盘上呈现大面积的红色（图 4.8(f)）。且可以观察到，缺陷情况的图片是在正常情况下发生不同程度的形变造成的。

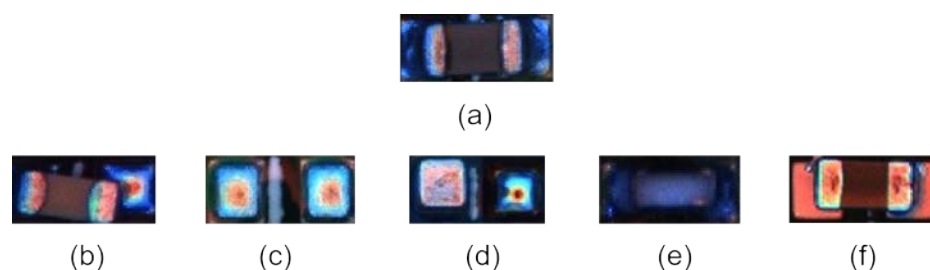


图 4.8 单个元件缺陷图

(2) 数据预处理与扩充

由于本文的 AOI 系统在采集图像时直接暴露在室内的光照环境下，在环境干扰的因素上缺少考虑，而焊锡膏对光照非常敏感，不同程度的光照会直接影响到锡膏反射出来的颜色。因此，有必要对采集数据进行预处理。本文首先对图像进行了滤波处理，以抑制由光照等环境因素造成的脉冲噪声和高斯噪声；其次，为了进一步增强锡膏的颜色特征，使用局部直方图均衡化增加图像的对比度；另外，由于本文模型固定了输入图像尺寸，且该尺寸远小于原始采集图的大小，在训练过程中采用有重叠的裁剪方法，将原图切分成多个较小图，处理后的图像尺寸为 1200x1200，在缩放后能够较好地适应模型输入尺寸，且同时扩充了数据量。

然而，工业缺陷数据集的稀少性不在于采集图像的数量，而在于有缺陷图像的数量。大多数情况下，工业场景中的数据存在着极大的不平衡，绝大多数都是正常样本，而不合格的样本只占据极少的比例，这非常不利于模型的训练。因此，要解决数据问题，就必须从缺陷样本本身着手。考虑到 PCBA 主板本身的特征较为单一，复杂的是贴装在主板上的元件，而前文中提到，我们在工厂中收集到的另一种数据是真实的单一原件缺陷图，且数据量充足，已经满足了缺陷本身多样性的要求。对该数据的利用极为重要，由于单一元件图中还存在着较多的冗余背景面积，必须先将元件轮廓区域，亦即感兴趣区域（Region of Interest, RoI）

提取出来，如图 4.9 所示。

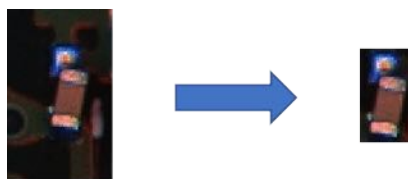


图 4.9 单一元件图 RoI 提取

这样得到的 RoI 图像就与目标检测任务中的预测框图类似，在此基础上，本文提出一种 PasteChips 方法，来生成大量的可标注数据，具体可分为以下两步。第一，利用工具软件 PhotoShop 将提取到的 RoI 图像直接粘贴到拍摄的 PCBA 局部图上，为了增强方法的鲁棒性，在粘贴时先将 RoI 调整到图中存在的某个正常元件的大小，并覆盖之，以伪造真实缺陷，这样就能最大程度地保留原图像的特征。第二，由于 PS 仍然需要手动处理，在需要大量数据时仍然显得捉襟见肘。因此，本文还使用了另一种更加智能化的增强方式，即先制作少量的无元件主板模板，利用算法直接将 RoI 随机粘贴到模板的任意位置，且 RoI 之间不允许重叠，该方法可以生成任意数量的标注数据，由于该算法生成的图像与真实图像差异较大，我们将其定义为伪标签数据，其流程见算法 4.1。生成的效果图如图 4.10 所示，其中 (a) 图是第一种方法得到，(b) 图由第二种方法得到。由于两种数据的差异性较明显，下文将分别进行不同的处理。

算法 4.1 伪标签数据生成算法

Input: RoI、模板
Output: 伪标签图像

- 1: **for** $i = 1$ to N **do**
- 2: 缺陷元件随机粘贴;
- 3: **while** 粘贴位置不冲突 **do**
- 4: 粘贴;
- 5: 记录粘贴位置坐标用于标注;
- 6: **end while**
- 7: 正常元件随机粘贴;
- 8: **while** 粘贴位置不冲突 **do**
- 9: 粘贴;
- 10: **end while**
- 11: **end for**

(3) 数据标注

经过上述的多种数据扩充方法，最终得到三种用于训练的图片共同组成最终的 SMTC 数据集，包括实际缺陷图，人工 PS 合成缺陷图和自动合成缺陷图。由于数据的形式不同，我们将前两者合并称之为 SMTC-Origin，而将自动合成的缺陷图视为伪标签数据集，称之为 SMTC-Pseudo。SMTC-Origin 包括 1189 张静态图片，其中 929 张用于训练，130 张用于验证，130 张用于测试，特别地，测试集全部采用实际缺陷图。SMTC-Pseudo 包括 5000 张静态图片，全部用于训练。

上面所述的数据集，获取的方式不同，处理的方式也各有差异。对于 SMTC-

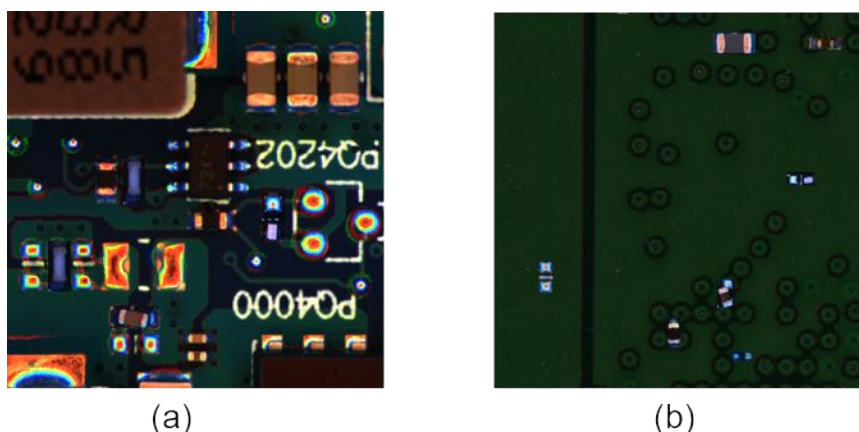


图 4.10 PasteChips 两种方法生成的图像

Pseudo, 在生成的时候会自动记录元件粘贴的位置坐标, 从而自动标注。SMTC-Origin 是需要手动标注的数据, 根据五种缺陷类别 (偏移、缺件、立碑、侧立和少锡), 使用 labeling 标注软件对数据集进行手动标注。

2. DeepPCB

DeepPCB 数据集是一个 PCB 裸板缺陷数据集, 它由上海交通大学的团队收集, 图片全部由线性 CCD 相机扫描得到。特别地, 为了消除光照干扰, 采用了二值化的图像处理方式。该数据集一共包含 1500 张静态图片, 每张图像的尺寸为 640x640, 其中训练集有 1050 张, 验证集有 150 张, 测试集有 300 张。对于 6 种常见的 PCB 裸板缺陷, 包括开路、短路、咬伤、杂散、针孔和伪铜做出了标注。尽管 PCB 裸板缺陷和本文的 SMT 贴片元件缺陷不同, 但同属于工业上 PCB 的缺陷问题, 且两者都存在着不规则的形变目标和小目标, 因此将 DeepPCB 作为实验数据集。图 4.11 展示了部分 DeepPCB 数据集的示例。

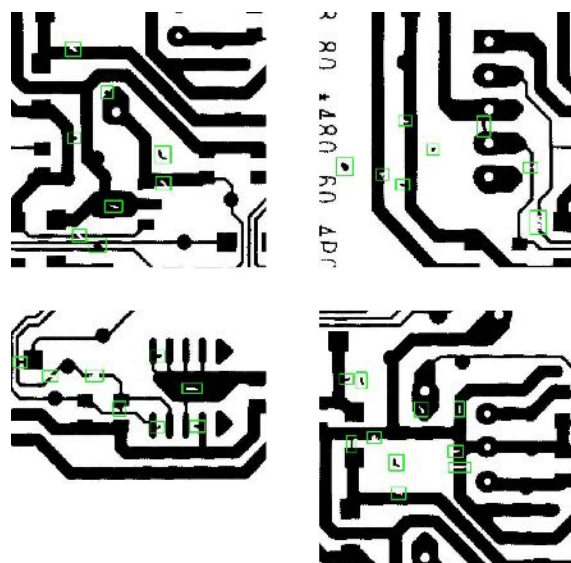


图 4.11 DeepPCB 数据集部分示例

3. VisDrone2019

VisDrone2019 数据集的收集工作是由位于天津大学 AISKYEYE 团队负责完成的。这个数据集包括了 10209 张通过无人机拍摄的静态图片，分配情况如下：其中 6471 张图片被归入训练集，548 张图片用作验证集，而 3190 张图片则被设置为测试集。数据集中包含的对象标注种类共计 10 类，涵盖行人、成人、汽车、面包车、巴士、卡车、摩托车、自行车、遮阳篷三轮车和三轮车。该数据集存在的难点在于检测对象数量巨大且有聚集现象、部分目标过小、不同的数据分布，部分示例如图 4.12 所示。由于 SMTc 数据集同样存在着目标小、物体密度大等，恰好与 VisDrone2019 的特点相似，因此本文将 VisDrone2019 作为实验数据集。



图 4.12 VisDrone2019 数据集部分示例

4.4.2 实验环境

本文算法全部在同一个环境中进行实验，软硬件的关键配置参数如下表所示。

表 4.1 检测实验环境配置参数

类型	型号	参数
操作系统	Ubuntu	20.04
GPU	NVIDIA Geforce RTX3090	24GB
Python	-	3.8.0
CUDA	-	11.6
Pytorch	-	1.11.0

4.4.3 评价指标

目标检测模型常用的指标包括精确率 (Precision, P), 召回率 (Recall, R), 均值平均精度 (mean Average Precision, mAP), GFLOPS。另外, 本文作为一种轻

量级的工业缺陷检测模型，对实时性的评估也尤为重要，本文还采用 FPS 作为评估模型的关键指标。下面将分别对各项指标作详细介绍。

为了理解以上指标，首先引入混淆矩阵的概念，如表 4.2 所示，将模型预测的结果分为四种类别：TP，表示真正类；FP，表示假正类；FN，表示假负类；TN，表示真负类。

表 4.2 混淆矩阵

	实际为正	实际为负
预测为真	TP(True Positive)	FP(False Positive)
预测为负	FN(False Negative)	TN(True Negative)

精确率 P 表示所有预测出目标中检测正确的概率，即被预测为正样本中实际为正类的比例，计算公式如下：

$$Precision = \frac{TP}{TP + FP} \quad (4.18)$$

召回率 R 表示目标被检出的概率，即所有实际为正类的样本中被预测为正确的比例，计算公式如下：

$$Recall = \frac{TP}{TP + FN} \quad (4.19)$$

在设定分类器置信度阈值（取值范围为 $[0, 1]$ ）之后，根据每个目标预测概率的大小，所有比阈值大的都为正例，比阈值小的都为负例。当阈值从 0 到 1 变化时，对应的 Precision 和 Recall 也不同，且两者相互制约。将每个 (Precision, Recall) 连接起来，即可绘制出 PR 曲线。不难得出，当两者同时达到最大时，模型的性能最优。因此对应的，PR 曲线与坐标轴的所包围的面积最大时效果最好，该面积的数值即是平均精确率 (Average Precision, AP)。每个类别都会对应一个 AP，将所有类别的 AP 之和取均值即为精度 (mean Average Precision, mAP)，其范围为 $[0, 1]$ 。在此基础上，本节中还采用 mAP50 和 mAP50-95，分别代表 IoU 阈值为 0.5 时的 mAP 和从 0.5 到 0.95 范围内步长为 0.05 的多个 mAP 的平均值。

GFLOPS (Giga Floating-point Operations Per Second) 表示每秒 10 亿次的浮点运算数，可以用来衡量模型的就算复杂度。

FPS (Frame Per Second) 表示模型每秒可以处理的图像数量，是对评价模型实时性最关键的指标。

4.4.4 实验结果分析

1. 对比实验

为了验证本文方法的有效性，我们将 SMT-DETR 与其他常见的实时目标检测模型在这两个数据集上进行了对比实验。由于所有 YOLO 算法都需要耗时的

NMS 后处理操作并计算最终检测时间，因此本文以 FPS 为衡量指标，从实时性角度出发，选取了 FPS 相当的 yolov5s、yolov6s、yolov7-tiny、yolov8s 和 RT-DETR 进行对比实验。所有模型训练时的部分超参数保持一致，如表 4.3 所示：

表 4.3 模型训练时部分超参数设置

超参数	参数信息
epoch	200
batchsize	4
初始学习率	0.0001
权重衰减	0.0001
优化器	AdamW
动量	0.9

在 SMTC-Origin 上得到的结果见表 4.4。可以观察到，相较于目前的先进模型而言，SMT-DETR 在性能上都有着不错的提升。与经典的实时检测器 yolov5、yolov6 和 yolov7 相比时，SMT-DETR 的 mAP 平均提升了 6 个百分点左右，并且 FPS 也是最高；与 yolo 系列中最先进的 yolov8 相比时，本文方法不仅做到了检测精度和速度同时超越，甚至在模型的尺度上也更为轻量化；与基准模型 RT-DETR 相比时，在保持 FPS 相当的同时，以超出 4.8% 的精度大幅领先。图 4.13 展示了 RT-DETR 和本文方法在 SMTC 数据集上的检测效果，对于相同的测试样本，其他比较模型都存在不同程度的误报问题，而只有 SMT-DETR 准确地检测出了所有缺陷。

表 4.4 SMTC-Origin 数据集对比实验结果

model	GFLOPS	P(%)	R(%)	mAP50(%)	mAP50-95(%)	FPS
yolov5s	15.8	82.8	77.2	80.7	38.8	52
yolov6s	44.0	84.3	72.8	80.2	38.9	86
yolov7-tiny	13.1	83.6	78.6	82.5	39.3	32
yolov8s	28.4	84.0	76.5	82.6	41.2	84
rt-detr-r18	57.0	82.8	75.5	79.9	40.1	90
ours	53.4	87.6	83.5	87.4	47.2	92

在 DeepPCB 数据集上得到的结果如表 4.5 所示。由于数据在输入不同模型时被调整为相同大小，因此此处省略了 GFLOPS 和 FPS。我们的方法实现了 98.6% 的 mAP，与目前的 SOTA 模型 (98.6% mAP @ 62 FPS) 持平，特别地，SMT-DETR 拥有更快的检测速度，以 92 的 FPS 领先。相较于表中的实时检测模型，不管在精度还是速度上都实现了超越。

在 VisDrone-DET2019 上得到的结果如表 4.6 所示。可以看出，本文方法实现了 46.8% 的 mAP，优于这些实时检测器。此外，精度在 VisDrone-DET2019 数据集的榜单上排名第三，仅次于 PP-YOLOE-plus^[68] 和 CZ Det^[69]。不过，就模型规模而言，前两位的在模型尺度上都比 SMT-DETR 大得多。

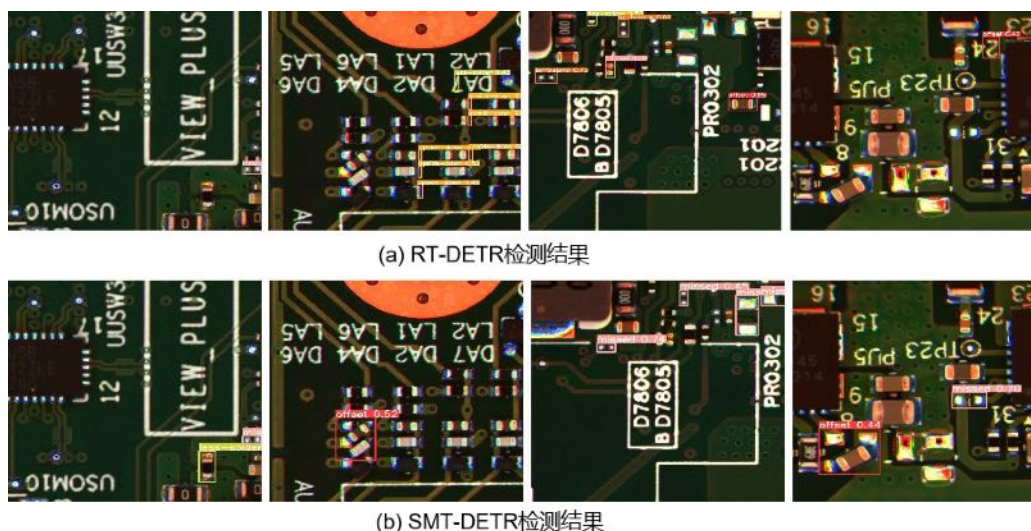


图 4.13 不同模型在相同的 SMT-Origin 数据集测试样本上的检测结果

表 4.5 DeepPCB 数据集对比实验结果

model	P(%)	R(%)	mAP50(%)	mAP50-95(%)
yolov5s	92.1	88.8	94.6	63.5
yolov6s	90.6	87.5	95.2	63.2
yolov7-tiny	97.3	93.6	97.3	51.8
yolov8s	90.4	83.9	92.9	60.3
rt-detr-r18	99.0	96.9	98.5	65.8
ours	98.8	97.3	98.6	67.3

表 4.6 VisDrone-DET2019 数据集对比实验结果

model	P(%)	R(%)	mAP50(%)	mAP50-95(%)
yolov5s	44.5	31.7	31.7	17.5
yolov6s	46.5	35.6	35.8	21.4
yolov7-tiny	48.5	41.7	40.2	15.6
yolov8s	50.8	38.2	38.8	23.4
rt-detr-r18	59.4	43.8	45.0	27.4
ours	60.8	44.7	46.8	28.7

2. 消融实验

为了验证本文提出的改进模块的有效性，在本节中，进行了一系列的消融实验。我们以 RT-DETR 为 baseline，将提出的改进模块依次加入模型进行实验，包括 ACBlock，LPE 和 IMIoU。图 4.14 显示了模型训练过程中，加入不同模块后的 mAP 变化曲线。可以观察到，添加的模块都有助于最终精度的提高。定量的比较数值在表 4.7 中展示。

(1) ACBlock: 表 4.8 显示了 ACBlock 插入在主干网络的不同位置时的效果比较。将 ACBlock 插入到 S7 时，检测器的精度达到了最佳。表 4.7 的第 4 行显示了模型在加入 ACBlock 后，除了 FPS 略微下降外，其余指标全部都有提升。DCNv2 相较于标准卷积，需要更多的参数量参与计算，因此 FPS 降低是在预料

表 4.7 改进模块消融实验

	P(%)	R(%)	mAP50(%)	mAP50-95(%)	FPS
RT-DETR(baseline)	84.8	81.1	84.3	43.1	91
+ACBlock	86.9	83.2	85.5	44.5	87
+ACBlock+LPE	87.7	82.6	86.2	45.3	89
+ACBlock+LPE+IMIou	87.6	83.5	87.4	47.2	92

表 4.8 将 ACBlock 插入到主干网络的不同位置的结果。{S5,S6,S7} 表示同时插入到 S5,S6,S7 三个位置, {S5} 表示只插入到 S5, 其他同理。

	P(%)	R(%)	mAP50(%)	mAP50-95(%)	FPS
{S5,S6,S7}	84.9	82.2	84.4	44.5	81
{S5}	87.4	83.3	86.1	44.1	89
{S6}	85.1	82.4	84.9	44.6	88
{S7}	87.6	83.5	87.4	47.2	92

之中的, 不过在 OrthoAttention 的帮助下, 下降的幅度很小。总体来看, ACBlock 带来的精度提升是完全值得的。

(2) LPE: 将原本的正余弦绝对位置编码改造为 LearnedPositionalEncoding 后, mAP 也得到了提升。

(3) IMIoU: 从图 4.14 中的训练过程可以看出, 加入 IMIoU, 也就是本文的最终模型 SMT-DETR 后, mAP 曲线收敛更快, 精度更高。表三最后一行显示了最终模型的准确率, 与基线相比, SMT-DETR 的 mAP50 和 mAP50-95 分别提高了 3.1% 和 4.1%。此外, 我们还与 CIoU、DIoU、EIoU^[70] 和 SIoU^[71] 进行了比较, 结果如图 4.15 所示, 可以看出 IMIoU 在训练过程中的损失值曲线优于其他算法。大量实验验证了 IMIoU 的有效性, 它不仅对小目标的检测性能更好, 而且收敛速度更快。

3. SMTC-Pseudo 实验

考虑到真实可标注数据数量较少, 可能导致模型训练不够充分而引起欠拟合的情况, 且大量的单一元件缺陷图未得到充分利用。本文采用 PasteChips 方法生成了 5000 张静态数据用于扩增数据, 并称之为 SMTC-Pseudo 数据集。在本章实验中, 验证集与测试集保持不变, 新增的 SMTC-Pseudo 数据集全部用于模型的训练过程。结果如表 4.9 所示。

表 4.9 SMTC-Pseudo 实验结果

P(%)	R(%)	mAP50(%)	mAP50-95(%)	FPS
89.0	87.0	88.3	46.5	90

实验结果表明, 经过 SMTC-Pseudo 扩展数据训练后, 模型的 mAP50 相较于之前提升了 0.6%, 且 P、R 也均有提升, 但 mAP50-95 略微下降。总体而言, 经过数据扩增后的模型有着提升, 这说明 SMTC-Pseudo 数据能够在一定程度上帮助网络更加充分地学习到不同贴片元件缺陷的特征。

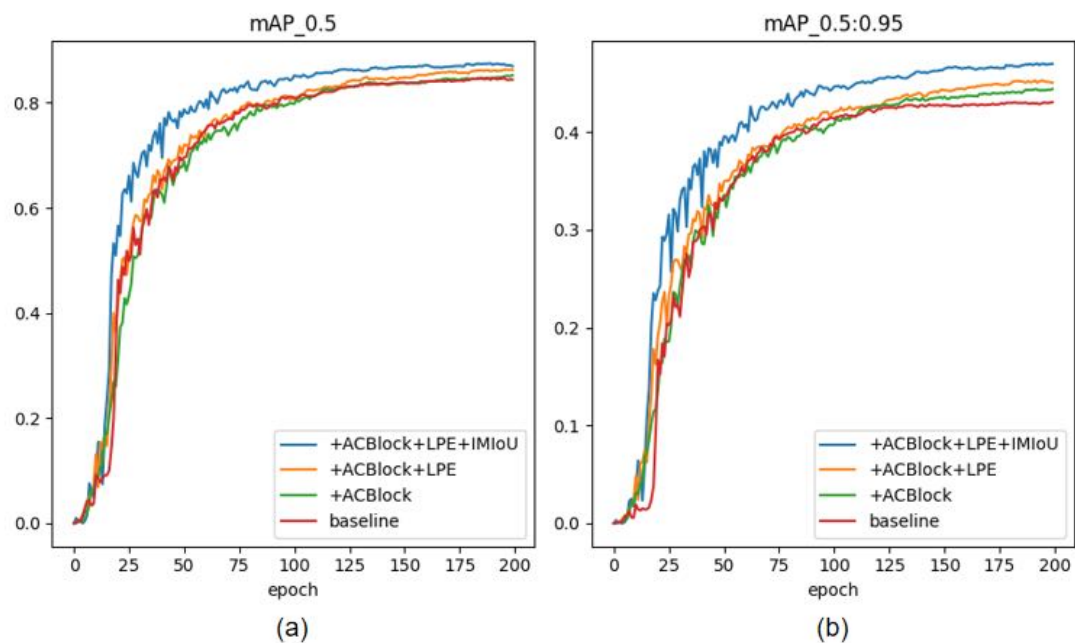


图 4.14 模型训练时 mAP 曲线. (a) mAP50; (b) mAP50-95.

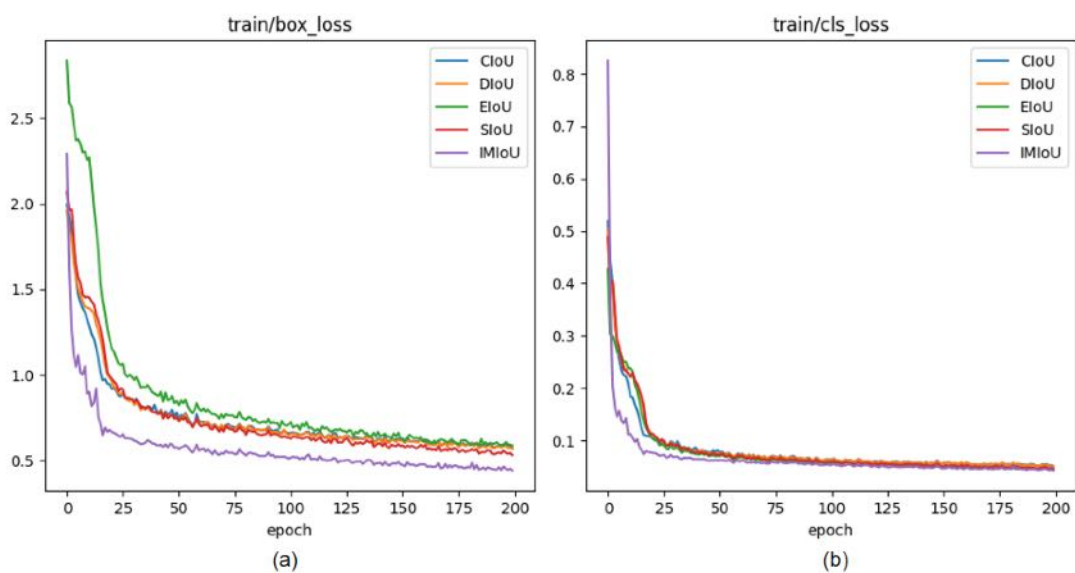


图 4.15 模型训练时 loss 曲线 (a) L_{box} . (b) L_{cls} .

4.5 本章小结

本章针对 SMT 产线中贴片元件经过回流焊后的缺陷检测问题，提出一种基于自适应协同特征的 DETR 缺陷检测算法，实现了多种复杂缺陷元件的有效检测。具体来说，通过实际采集和数据增强方法建立起 SMT 贴片元件缺陷数据集，以应对相关数据稀少的情况。在模型设计上，采用轻量化的 RT-DETR 作为基准模型，将 transformer 强大的建模能力应用于实时检测领域；利用可变形卷积和通道注意力机制设计处自适应协同模块以替代部分传统的卷积模块，能够更好地关注到元件缺陷的形变特征；提出一种全新的 IMIoU 损失函数，对 PCB 图像中的微小贴片元件更加友好，并且具备较快的收敛速度。在 SMTC 数据集和公开数据集（DeepPCB 和 VisDrone2019）上的实验表明，本文方法在速度和精度上取得了良好的平衡，并且具有较好的泛化能力。未来的研究将在保证实时性的情况下提升模型的检测精度，使模型更好地适用于工业实时检测任务。

第 5 章 总结与展望

5.1 现阶段工作总结

SMT 技术的出现给电子制造行业带来了巨大的便利，它通过直接将无引脚或短引线的表面组装元器件安装在印制电路板的表面，使得元件封装的流程完全由机器自动化实现，这极大地提升了生产效率。随着技术的不断成熟，PCB 表面的贴装元件越来越密集，对元件在经过回流焊后的缺陷检测问题提出了更高的挑战。当前主流的方法是采用基于图像处理的 AOI 技术实现对异常产品的自动检测。然而，现有的 AOI 技术仍存在较大的不足，主要表现出响应速度缓慢、缺陷误检率高、投入成本大等问题。本文总结出两个引起当前问题的原因。一方面，问题源于图像拼接。拍摄 PCBA 图像时，相机视场和图像分辨率存在双重要求，需要拍摄多张局部图像并进行拼接以形成完整图像，随后再与标准模板进行比对。拼接算法在这一过程中起着至关重要的作用。另一方面，问题在于缺陷检测。采用模板匹配的缺陷检测方法存在着诸多弊端，因为模板和拍摄图像都容易受到外在干扰，并且缺陷判定规则单一，导致误检率居高不下。为应对机器判断效率疲软的情况，又需要设置人工复判环节，导致成本上升。

本文从引起上述两个问题的根本原因出发，分别对为两个环节提供支持的关键技术进行改进。对于图像拼接问题，提出一种基于 SURF 主要距离的主板图像拼接算法，实现了高清 PCB 图像的高效拼接。对于缺陷检测问题，提出一种基于自适应协同特征的 DETR 缺陷检测算法，在本文从某消费电子制造商的 SMT 产线上收集的 SMT 贴片元件数据集上的实验结果表面，该方法具有较好的效果。基于上述研究，本文实现了 AOI 系统中关键技术上的改进，这些算法均将部署在课题组搭建的 AOI 设备上，预计建成后将有效提升制造商的整体生产效率，并产生巨大经济效益。

本文主要研究内容和创新点总结如下：

(1) 针对 AOI 系统采集的多张 PCB 局部图像拼接成整板图像的问题，通过将硬拼接和软拼接方法相结合，提出了一种基于 SURF 主要距离的主板图像拼接算法。这一算法实现了 PCB 图像的实时高质量拼接。同时，通过提出基于划分区域的特征点提取方法和基于主要范围的匹配对细筛选方法，对 SURF 算法进行改进，在算法速度和准确度上得到了一定的提升。在处理震动引起的不同图像偏移量变化情况时，利用改进后的 SURF 算法提取出待拼接图像的特征点并匹配，将匹配对之间的主要距离作为偏移量，实现每次拼接不同对象时的实时调整。为了解决光照不均导致的拼接缝色差问题，采用了基于加权像素的空域融合法。通过偏移补偿方法，保证所有列拼接图像尺寸与第一列图像一致。在 PCB

整板图像拼接实验中，本文方法能够稳定地实现图像拼接过程，且仅需约 16 秒的时间，同时保留图像高清晰度。最终结果显示，图像错位情况得到明显改善。

(2) 针对 SMT 产线中贴片元件经过回流焊后可能出现的缺陷，采用了深度学习技术，提出了基于自适应协同特征的 DETR 缺陷检测算法，有效检测多种复杂缺陷元件。为解决数据稀缺问题，构建了包含多种缺陷情况的 SMT 贴片元件数据集。在注重检测速度和精度的工业环境中，选择了轻量化的 RT-DETR 作为基准模型，该模型不仅拥有 transformer 的强大建模能力，而且保持了良好的实时性。设计了自适应协同模块来关注贴片元件特殊的缺陷形态，以代替一部分传统卷积模块，能够更精准地识别缺陷形变特征。此外，论文还提出了一种新的损失函数，它对 PCB 图像中的微小贴片元件检测更为友好，并能够加快模型的收敛速度。实验结果显示，该方法在 mAP50 指标上达到 88.3%，比基准方法提升了 8.4%，FPS 指标达到 90，在速度和精度上取得良好平衡，超越当前所有实时目标检测模型，并且在一些公开数据集上也具有不错的效果。

5.2 未来工作展望

本文对当前 AOI 系统中关键技术的改进，包括图像拼接技术和缺陷检测技术，对课题组搭建的 AOI 系统具有一定的促进意义，但仍存在着一些不足，需要继续改进，主要有以下几点：

(1) 提升拼接算法对不同场景的适应能力。当前实现的拼接算法高度依赖于硬件系统采集 PCB 图像的路线，并且其中的部分参数是根据既定的规则而精心设计的。在采集路线发生变化时，相应的参数也需要重新设定，过程稍显繁琐。

(2) 扩充数据集。目前针对 SMT 场景下的贴片元件缺陷检测问题还没有相关的公开数据集，导致对该问题的研究局限性较强。本文通过收集实际图像和数据增强的方法，初步构建了 SMTC 数据集，但真实数据量仍然较少，未来的工作将继续在实际产线上采集数据，并以此提升模型的性能。

(3) 对模型进行轻量化处理。本文研究的算法针对实际应用场景，需要在 AOI 系统的工控机上部署。由于工控机算力的限制，目前无法使用拥有更高精度但尺度更大模型进行检测，未来将地关注于知识蒸馏、剪枝等模型压缩技术，使得部署的模型性能更佳。

(4) 增加电路检测方法。目前所采用的基于视觉检测的模型只能做到 PCB 的外观检测，无法对电路的缺陷进行检测，例如虚焊和假焊等不良情况，这仍然依赖于 AOI 质检后续的功能测试环节。

参 考 文 献

- [1] 胡跃明, AN C, JUAN D, 等. SMT 装备业的发展机遇与挑战[C]//第九届中国高端 SMT 学术会议论文集. 2015: 40-52.
- [2] 况延香, 朱颂春. 微电子封装技术对 SMT 的促进作用[J]. 电子与封装, 2004: 1-5.
- [3] 吴一全, 赵朗月, 苑玉彬, 等. 基于机器视觉的 PCB 缺陷检测算法研究现状及展望[J]. 仪器仪表学报, 2022, 43(1-17).
- [4] 鲁华杰, 张建, 王亚宁, 等. 自动测试技术和设备现状及发展趋势[J]. 海峡科技与产业, 2017: 63-64+67.
- [5] 茅海峰. SMT 检测中 AOI 技术的应用[J]. 科技信息, 2010: 57+47.
- [6] 李海玉, 王雷, 杨晓燕, 等. SMT 生产中的 AOI 检测技术现状分析[J]. 时代农机, 2019, 46: 54-55.
- [7] 王玉萍, 郭峰林. 基于机器视觉的 PCB 板缺陷检测系统关键技术研究[J]. 科技通报, 2017, 33: 101-105.
- [8] 罗运, 贺翔, 丁诗婕. 图像拼接算法发展综述[J]. 现代计算机, 2021(08): 78-82.
- [9] SZELISKI R, et al. Image alignment and stitching: A tutorial[J]. Foundations and Trends® in Computer Graphics and Vision, 2007, 2(1): 1-104.
- [10] BROWN M, LOWE D G. Invariant features from interest point groups.[C]//Bmvc: volume 4. 2002: 398-410.
- [11] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [12] BAY H, TUYTELAARS T, GOOL L V. Surf: Speeded up robust features[C]//Proceedings of the 9th European conference on Computer Vision - Volume Part I. 2006.
- [13] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05): volume 1. Ieee, 2005: 886-893.
- [14] LI X B, BAO L, WANG S G, et al. Study on image mosaic and registration of electronic component through surface mounting technology[J]. Advanced Materials Research, 2012, 345: 113-118.
- [15] 裴新权, 周亚丽, 张奇志, 等. 一种应用于 PCB 裸板质量检测的图像拼接算法[J]. 实验室研究与探索, 2017, 36(08): 24-28.
- [16] 杨帆, 金国强, 金一, 等. 基于 SIFT 特征提取的多重复单元 PCB 板图像拼接[J]. 工业控制计算机, 2018, 31(09): 3-4+7.
- [17] SUNDARAJ K. Pcb inspection for missing or misaligned components using background sub-

- traction[J]. WSEAS transactions on information science and applications, 2009, 6(5): 778-787.
- [18] CAI N, LIN J, YE Q, et al. A new ic solder joint inspection method for an automatic optical inspection system based on an improved visual background extraction algorithm[J]. IEEE Transactions on Components, Packaging and Manufacturing Technology, 2015, 6(1): 161-172.
- [19] ZHANG Z, WANG H, ZHOU S, et al. A cascaded zoom-in method for defect detection of solder joints[C]//2018 14th IEEE International Conference on Signal Processing (ICSP). IEEE, 2018: 1081-1086.
- [20] KIM Y G, LIM D U, RYU J H, et al. Smd defect classification by convolution neural network and pcb image transform[C]//2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS). IEEE, 2018: 180-183.
- [21] KIM Y G, PARK T H. Smt assembly inspection using dual-stream convolutional networks and two solder regions[J]. Applied Sciences, 2020, 10(13): 4598.
- [22] PANG G, SHEN C, CAO L, et al. Deep learning for anomaly detection: A review[J]. ACM computing surveys (CSUR), 2021, 54(2): 1-38.
- [23] LIU G, WEN H. Printed circuit board defect detection based on mobilenet-yolo-fast[J]. Journal of Electronic Imaging, 2021, 30(4): 043004-043004.
- [24] DING R, DAI L, LI G, et al. Tdd-net: a tiny defect detection network for printed circuit boards [J]. CAAI Transactions on Intelligence Technology, 2019, 4(2): 110-116.
- [25] 任斌, 程良伦. AOI 机器视觉系统中检测光源的分析和设计[J]. 微计算机信息, 2009, 25(42-44+145).
- [26] 张威. 空间域中均值滤波与中值滤波去噪的应用研究[J]. 产业科技创新, 2020: 65-66.
- [27] 沈学利, 王肃国. 基于 FPGA 的改进型中值滤波算法研究[J]. 微电子学与计算机, 2014, 31(1): 4.
- [28] 李惠芬, 蒋向前, 李柱. 高斯滤波稳健性能的研究与改进[J]. 仪器仪表学报, 2004, 25(5): 5.
- [29] 孙抗, 汪渤, 周志强, 等. 基于双边滤波的实时图像去雾技术研究[J]. 北京理工大学学报, 2011, 31(7): 5.
- [30] 邵平, 杨路明. 基于模板分解和积分图像的快速 Kirsch 边缘检测[J]. 自动化学报, 2007, 33(8): 6.
- [31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [32] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
- [33] GIRSHICK R. Fast r-cnn[C/OL]//2015 IEEE International Conference on Computer Vision

- (ICCV). 2015: 1440-1448. DOI: 10.1109/ICCV.2015.169.
- [34] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [35] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
- [36] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer, 2016: 21-37.
- [37] REDMON J, FARHADI A. Yolo9000: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
- [38] REDMON J, FARHADI A. Yolo3: An incremental improvement[A]. 2018.
- [39] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolo4: Optimal speed and accuracy of object detection[A]. 2020.
- [40] LI C, LI L, JIANG H, et al. Yolo6: A single-stage object detection framework for industrial applications[A]. 2022.
- [41] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Yolo7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.
- [42] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[C]//International Conference on Learning Representations. 2021.
- [43] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]//European conference on computer vision. Springer, 2020: 213-229.
- [44] MILLS-TETTEY A, STENT A, DIAS M B. The dynamic hungarian algorithm for the assignment problem with changing costs[J]. carnegie mellon university, 2007.
- [45] ZHU X, SU W, LU L, et al. Deformable detr: Deformable transformers for end-to-end object detection[C]//International Conference on Learning Representations. 2021.
- [46] LIU S, LI F, ZHANG H, et al. DAB-DETR: Dynamic anchor boxes are better queries for DETR[C]//International Conference on Learning Representations. 2022.
- [47] LI F, ZHANG H, LIU S, et al. Dn-detr: Accelerate detr training by introducing query denoising [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 13619-13627.
- [48] ZHANG H, LI F, LIU S, et al. DINO: DETR with improved denoising anchor boxes for end-to-

- end object detection[C]//The Eleventh International Conference on Learning Representations. 2023.
- [49] 王娟, 师军, 吴宪祥. 图像拼接技术综述[J]. 计算机应用研究, 2008(1940-1943+1947).
- [50] LEUTENEGGER S, CHLI M, SIEGWART R Y. Brisk: Binary robust invariant scalable key-points[C]//International Conference on Computer Vision. 2011.
- [51] RUBLEE E, RABAUD V, KONOLIGE K, et al. Orb: An efficient alternative to sift or surf [C]//2011 International Conference on Computer Vision. 2011: 2564-2571.
- [52] 索春宝, 杨东清, 刘云鹏. 多种角度比较 SIFT、SURF、BRISK、ORB、FREAK 算法[J]. 北京测绘, 2014(23-26+22).
- [53] ABEYWICKRAMA T, CHEEMA M A, TANIAR D. k-nearest neighbors on road networks: A journey in experimentation and in-memory implementation[J]. Proceedings of the VLDB Endowment, 2016, 9(6).
- [54] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395.
- [55] YANG Y, YANG M, HUANG S, et al. Multifocus image fusion based on extreme learning machine and human visual system[J]. IEEE Access, 2017: 6989-7000.
- [56] ZHANG H, XU C, ZHANG S. Inner-iou: More effective intersection over union loss with auxiliary bounding box[A]. 2023.
- [57] SILIANG M, YONG X. Mpdio: A loss for efficient and accurate bounding box regression [A]. 2023. arXiv: 2307.07662.
- [58] SALMAN H, PARKS C, SWAN M, et al. Orthonets: Orthogonal channel attention networks [C]//2023 IEEE International Conference on Big Data (BigData). IEEE, 2023: 829-837.
- [59] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [60] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 764-773.
- [61] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 9308-9316.
- [62] REZATOFIGHI H, TSOIN, GWAK J, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 658-666.
- [63] ZHENG Z, WANG P, LIU W, et al. Distance-iou loss: Faster and better learning for bounding box regression[C]//Proceedings of the AAAI conference on artificial intelligence: volume 34.

- 2020: 12993-13000.
- [64] ZHANG J, DING R, BAN M, et al. Fdsnet: An accurate real-time surface defect segmentation network[C]//ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022: 3803-3807.
- [65] JESSURUN N, DIZON-PARADIS O P, HARRISON J, et al. Fpic: A novel semantic dataset for optical pcb assurance[J]. ACM Journal on Emerging Technologies in Computing Systems, 2023, 19(2): 1-21.
- [66] TANG S, HE F, HUANG X, et al. Online pcb defect detector on a new pcb defect dataset: abs/1902.06197[A/OL]. 2019. <https://api.semanticscholar.org/CorpusID:62841418>.
- [67] DU D, ZHU P, WEN L, et al. Visdrone-det2019: The vision meets drone object detection in image challenge results[C]//Proceedings of the IEEE/CVF international conference on computer vision workshops. 2019: 0-0.
- [68] LONG X, DENG K, WANG G, et al. Pp-yolo: An effective and efficient implementation of object detector[A]. 2020. arXiv: 2007.12099.
- [69] MEETHAL A, GRANGER E, PEDERSOLI M. Cascaded zoom-in detector for high resolution aerial images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 2045-2054.
- [70] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient iou loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [71] GEVORGYAN Z. Siou loss: More powerful learning for bounding box regression[A]. 2022. arXiv: 2205.12740.

致 谢

草长莺飞二月天，拂堤杨柳醉春烟，如今已是我在科大的第三个春天，三年的硕士生涯也即将步入尾声。在毕业论文完成之际，我不禁感慨良多，良师益友是我在读研期间莫大的收获，在此我向所有在学习和生活中帮助过我的老师和同学们给予深刻感激之情。

感谢我的两位导师赵云波教授和许镇义副研究员。感谢赵老师给了我读研的机会，使我能够来到科大，在这里我学习到了很多的专业知识，也结交了诸多的朋友。在与老师的交流中，我学习到的科研思维和做事准则，都是我人生中的宝贵财富，在今后的人生中也将继续受益于此。感谢许老师对于课题上的悉心指导，在科研中遇到的诸多问题都能一一向老师请教。老师不仅是我科研路上的引导者，也是生活中亲切的朋友。

感谢实验室的同门，在学习生活中都给与我很多的帮助，很幸运能遇见你们。特别地，对于同课题组的小伙伴们，我们为了一个同样的科研目标一起奋斗了三年时光，彼此陪伴与帮助。感谢我的室友们，我们一起创造了 208 和谐的宿舍环境，三年朝夕相处，此后情谊永存。

感谢我的亲人朋友，特别是我的父母，多年来一如既往地支持我的学业，为我解决了学习之外所有的后顾之忧，成为你们的孩子是我莫大的幸运。希望父母能够身体健康、幸福生活。

最后，衷心感谢在百忙之中抽出时间评阅我的论文和出席论文答辩的各位专家、教授，你们辛苦了！

在读期间发表的学术论文与取得的研究成果

已发表会议论文

1. A Real-time Detection Method for SMT Chip Component Defects Based on Adaptive Collaborative Feature. Yunbo Zhao, Wangyou Gui, Yu Kang, Kehao Shi, Lijun Zhao, and Zhenyi Xu. In 2024 International Conference on Guidance, Navigation and Control (ICGNC 2024). (已录用)

专利

1. 赵云波, 桂旺友, 许镇义, 史珂豪, 王子曜. 贴片元件缺陷检测模型的训练方法、检测方法及其装置:CN202410107489.6[P].2024-05-17.
2. 许镇义, 桂旺友, 曹洋, 康宇, 赵云波. 用于 PCB 微小缺陷检测的单帧目标检测方法及其存储介质:CN202210453360.1[P].2022-07-12.